

# **Image Analysis for Segmentation of Psoriasis Lesion**

by

**Chong Chi Hung**

Dissertation submitted in partial fulfilment of  
the requirements for the  
Bachelor of Engineering (Hons)  
(Electrical and Electronic Engineering)

SEPTEMBER 2011

Universiti Teknologi PETRONAS  
Bandar Seri Iskandar  
31750 Tronoh  
Perak Darul Ridzuan

# CERTIFICATION OF APPROVAL

## **Image Analysis for Segmentation of Psoriasis Lesion**

by

**Chong Chi Hung**

A project dissertation submitted to the

**Electrical and Electronic Engineering Programme**


**Universiti Teknologi PETRONAS**

in partial fulfilment of the requirement for the

**BACHELOR OF ENGINEERING (Hons)**

**(ELECTRICAL AND ELECTRONIC ENGINEERING)**

Approved by,

  
A handwritten signature in black ink, appearing to read 'Ahmad Fadzil Mohamad Hani', is written over a horizontal line.

(Prof Ir Dr Ahmad Fadzil Mohamad Hani)

**UNIVERSITI TEKNOLOGI PETRONAS**

**TRONOH, PERAK**

**September 2011**

## CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



---

CHONG CHI HUNG

## ABSTRACT

Psoriasis, a hereditary inflammatory skin condition that currently affects 2 – 3 % of the world's population, is marked by reddish, scaly rashes or lesions covered with scabs of dead skin. This dermatosis is currently not curable but the symptoms can be effectively controlled through an accurate assessment scheme and well-integrated medical care therapy. The National Psoriasis Foundation Medical Board has published a guideline that categorizes the severity of psoriasis – mild, moderate and severe, each characterized by the percentage of lesions on an individual's body surface area. However the caveat remains that the distinction between the different categories of severity is largely influenced by the clinical practitioner's subjectivity. As a result, PASI scoring is introduced. PASI (Psoriasis Area and Severity Index) is currently the gold standard method to measure psoriasis severity by evaluating the area, erythema, scaliness and thickness of the lesions. These 4 parameters require the lesions to be first segmented from the skin patches before they can be assessed and scored individually. This report thus investigates digital image analysis techniques to segment psoriasis lesions. In this work, 90 patients are categorized into groups of differing skin tones based on mean values in the  $L^*$  component. The 1000 new colours obtained through clustering of pixel values in the R, G and B component are used to construct three different skin lesion models. The validation of the three models is done by comparing the mean values of constructed models and original image, which are found to be the same. For segmentation of skin into involved and non-involved regions, iterative thresholding and Otsu's method are applied in 3 colour spaces, namely,  $I_1I_2I_3$ , CIE  $L^*a^*b^*$  and HSI. The average segmentation errors in the 3 colour spaces are then compared to select the best colour channel in which to perform the segmentation for either thresholding. The specificity and sensitivity analysis with the accompanying Type I error and Type II error are conducted as well. From the segmentation results of skin lesion models, it is found that segmentation in the  $I_3$  colour channel (for fair skin tone),  $I_3$  and  $b$  colour channels (for middle skin tone) and  $I_2$ ,  $b$  and  $S$  colour channels (for dark skin tone) yields high accuracy. The same thresholding method in



corresponding colour channels is then applied on 20 real skin samples. The segmented images are compared with the reference images to measure the accuracy of the proposed lesion segmentation method in different colour channels. Out of 20 cases, the segmentation method achieved accuracies of higher than 95% for 19 cases. The lowest accuracy obtained is for a particular skin-lesion patch with accuracies of 92 - 93%. The lower accuracy is due to wrinkled skin areas which have been exposed to unequally-distributed light leading to misclassification as lesions. For each different skin tone, the overall accuracy results show that the proposed colour channels are appropriate and accurate to carry out the Otsu's method for segmentation of psoriasis lesions.

## **ACKNOWLEDGEMENTS**

Throughout the 9 months leading to the completion of my final year project, there are several figures who I would like to record my deepest appreciation to.

First and foremost, I would like to express my utmost gratitude to my supervisor, Professor Ir Dr Ahmad Fadzil Mohamad Hani for his continuous guidance and insights. Despite his duties as the Deputy Vice Chancellor and being in charge of several postgraduate students, Professor Fadzil has consistently taken the time out to provide solid input to improve my report and laboratory work.

I would also like to extend my gratitude to Mr Esa Prakasa, the postgraduate student in the Centre of Intelligent Signal & Imaging Research (CISR) who has acted as my mentor. The same level of appreciation is extended to all my colleagues in the CISR for their advice and support.

Additionally, I would like to thank the FYP Coordinator, Dr Zuhairi Baharudin for his efforts in coordinating the FYP proceedings. Last but not least, thank you to my family members and friends who have supported me throughout the past 9 months. What I have achieved in this Final Year Project would not have been possible without the commitment given by all the individuals and parties involved, directly or indirectly.

## TABLE OF CONTENTS

Title	Page
<b>List of Figures</b>	<b>i</b>
<b>List of Tables</b>	<b>iii</b>
<b>1.0 Introduction</b>	<b>1</b>
1.1 Background of Study	1
1.1.1 Causes of Psoriasis	2
1.1.2 Classification of Psoriasis (Signs and Symptoms)	4
1.1.3 Incidences of Psoriasis	6
1.2 Problem Statements	9
1.3 Objective	9
1.4 Scope of Study	
<b>2.0 Literature Review</b>	
2.1 Colour Spaces	10
2.1.1 CIE $L^* a^* b^*$	10
2.1.2 HSI	11
2.1.3 $I_1 I_2 I_3$	11
2.2 Global Thresholding	12
2.3 Related Works	13
<b>3.0 Methodology</b>	
3.1 Data Acquisition	15
3.2 Construction and Validation of Skin-and-Lesion Model	17
3.3 Selection of Segmentation Techniques	19
3.3.1 Transformation to CIE $L^* a^* b^*$ , HSI and $I_1 I_2 I_3$	19
3.3.2 Segmentation of Lesion (Pixel-Based Thresholding)	20
3.4 Application of Proposed Techniques on Real Skin-Lesion Samples	21
3.5 Manipulation of Size of Lesion	21
3.6 Tools and Equipments Used	22
<b>4.0 Results and Discussion</b>	
4.1 Data Acquisition	26
4.2 Construction and Validation of Skin-and-Lesion Model	26
4.2.1 Grouping of Patients based on Skin Tones	26
4.2.2 Construction of new colours through clustering in R,	

	G and B component	28
4.2.3	Fuzzy c-means to eliminate overlap of lesion and skin cluster	34
4.2.4	Validation of Skin-Lesion Models	37
4.3	Selection of Segmentation Techniques	41
4.3.1	Segmentation Techniques Performed on Skin-Lesion Models	41
4.3.2	Analysis of Segmentation Techniques	43
4.4	Application of Proposed Techniques on Real Skin-Lesion Samples	47
4.4.1	Calculation of Average Error Percentage	49
4.4.2	Analysis of Accuracy Results	60
5.0	Timeline for Completion of Final Year Project	65
6.0	Conclusions	
6.1	Discussion	66
6.2	Contribution & Recommendation	68
	Reference	70
	Appendix	73

## **LIST OF FIGURES**

<b>Figure 1</b>	Different Types of Psoriasis
<b>Figure 2</b>	Different Skin Tones and Corresponding Lesion Colours
<b>Figure 3</b>	Original Images and Transformation into Different Colour Spaces
<b>Figure 4</b>	Segmentation on a Male Mannequin
<b>Figure 5</b>	CIE L* a* b* Colour Space
<b>Figure 6</b>	Skin Model
<b>Figure 7</b>	Skin-and-Lesion-Model
<b>Figure 8</b>	HSI Colour Space
<b>Figure 9</b>	Tools Used
<b>Figure 10</b>	Flow Chart for Results & Discussion
<b>Figure 11</b>	Retrieval of Skin and Lesion Samples
<b>Figure 12</b>	L* channel histogram for 90 patients
<b>Figure 13</b>	RGB channel histograms for fair-skin patients
<b>Figure 14</b>	Constructed Skin Model
<b>Figure 15</b>	Zoom-in of Constructed Skin Model
<b>Figure 16</b>	Pixel with R = 3, G = 123 and B = 8
<b>Figure 17</b>	RGB Colour Spaces
<b>Figure 18</b>	Probability Density Function for Dark Skin Tone
<b>Figure 19</b>	Probability Density Function for Middle Skin Tone
<b>Figure 20</b>	Probability Density Function for Fair Skin Tone

- Figure 21** Final Constructed Skin-and-Lesion-Model
- Figure 22** Thresholding through Differences of Saturation for Skin-Lesion Images
- Figure 23** Real Skin-Lesion Patches (Patch 1 - Patch 20)
- Figure 24** Real Skin-Lesion Patches (Fair Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 1 - Patch 8)
- Figure 25** Real Skin-Lesion Patches (Middle Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 9 - Patch 14)
- Figure 26** Real Skin-Lesion Patches (Dark Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 15 - Patch 20)
- Figure 27** Skin-Lesion Patch 15
- Figure 28** Skin-Lesion Patch 16
- Figure 29** Timeline for Completion of Final Year Project

## LIST OF TABLES

<b>Table 1</b>	Distribution of psoriasis cases according to age and gender among outpatients attending dermatological Clinic in Hospital Tengku Ampuan Rahimah Klang, Selangor, Malaysia
<b>Table 2</b>	Distribution of psoriasis according to ethnic groups among outpatients attending Hospital Tengku Rahimah, Klang, Selangor, Malaysia
<b>Table 3</b>	Area of body commonly affected with psoriasis
<b>Table 4</b>	L* Component Threshold Values
<b>Table 5</b>	Mean Value of L* Component of Patients
<b>Table 6</b>	Categorization of 90 Patients
<b>Table 7</b>	Cluster Mean Value for Fair Normal Skin Group
<b>Table 8</b>	Cluster Mean Value for Middle Normal Skin Group
<b>Table 9</b>	Cluster Mean Value for Dark Normal Skin Group
<b>Table 10</b>	Cluster Mean Value for Fair Lesion Group
<b>Table 11</b>	Cluster Mean Value for Middle Lesion Group
<b>Table 12</b>	Cluster Mean Value for Dark Lesion Group
<b>Table 13</b>	Colour Number for Construction of Fair Skin-Lesion Model
<b>Table 14</b>	Colour Number for Construction of Middle Skin-Lesion Model
<b>Table 15</b>	Colour Number for Construction of Dark Skin-Lesion Model
<b>Table 16</b>	Validation of Skin-Lesion Model (R,G and B Mean Values)
<b>Table 17</b>	Make-up Colours for Fair Skin Lesion Model

<b>Table 18</b>	Make-up Colours for Middle Skin Lesion Model
<b>Table 19</b>	Make-up Colours for Dark Skin Lesion Model
<b>Table 20</b>	Percentage Error by Colour Channel (Fair Skin)
<b>Table 21</b>	Percentage Error by Colour Channel (Middle Skin)
<b>Table 22</b>	Percentage Error by Colour Channel (Dark Skin)
<b>Table 23</b>	Colour Channel with < 5 % Error (3 Different Skin-Lesion Models)
<b>Table 24</b>	Skin-Lesion Tones and Appropriate Colour Channels
<b>Table 25</b>	Mean L* Value and Skin Tone for Skin Lesion Patches
<b>Table 26</b>	Specificity for Segmentation of Skin and Lesion (Patch 1 - Patch 20)
<b>Table 27</b>	Type I Error (Patch 1 - Patch 20)
<b>Table 28</b>	Sensitivity for Segmentation of Skin and Lesion (Patch 1 - Patch 20)
<b>Table 29</b>	Type II Error (Patch 1 - Patch 20)
<b>Table 30</b>	Percentage of Accuracy for Segmentation of Skin and Lesion (Patch 1 - Patch 20)



# CHAPTER 1: INTRODUCTION

## 1.1 Background of Study

References to this skin disease were first found in writings by Greek physician Hippocrates but psoriasis was identified by Robert Willan, an English dermatologist as an independent disease. Psoriasis is a chronic immune-mediated disease of the skin that causes rapid skin cell reproduction due to faulty signals in the immune system. However the body does not slough off the excess skin and the accretion of skin layers on the surface of the skin form lesions. This result in red, dry patches of skin covered with flaky silver, white patches (scales) commonly found at the skin area of the elbows, knees and scalp.

According to the formal description of psoriasis as developed by the experts at the American Academy of Dermatology,

*"[psoriasis is] chronic skin disease that is classically characterized by thickened, red areas of skin covered with silvery scale."*

Psoriasis affects both genders and usually manifests itself for the first time between the ages of 15 and 25. A study by Farber and Nall however suggests that the average age of onset as 27.8 years as proven by 35 % of patients involved [1]. Psoriasis tends to appear without warning and has been shown to wax and wane in severity during an individual's life. The disease though not fatal causes discomfort, embarrassment or disfigurement for the sufferers. There has been no found cure to date but psoriasis can easily be controlled through regulated treatment in the long term.

### 1.1.1 Causes of Psoriasis

The cause of psoriasis has long been linked to the faulty genetic component that is hereditary. Two schools of thought have surfaced with regard to the genesis of the disease. The first pins it down to the epidermis and its keratinocytes for causing excessive growth of the skin tissues. The second opinion, propounded by a group of scientists, tracks the linkage to the T cells (a type of white blood cell for cell-mediated immunity) for mistakenly triggering a reaction in the skin cells. The known fact however, is that certain medicines, including beta blockers, antimalarial drug

chloroquine etc, excessive alcohol consumption, cream-and-lotion-based-products can aggravate the disease. The disease is however not contagious.

A study carried out on people and their relatives with psoriasis shows a genetic or familial disposition to the disease [2]. This seems to suggest that psoriasis is purely a hereditary disease but scientists have also learned that not everyone who inherits genes for psoriasis gets psoriasis. It is contingent on the 'right' mix of genes and the existence or prolongment of an environmental trigger.

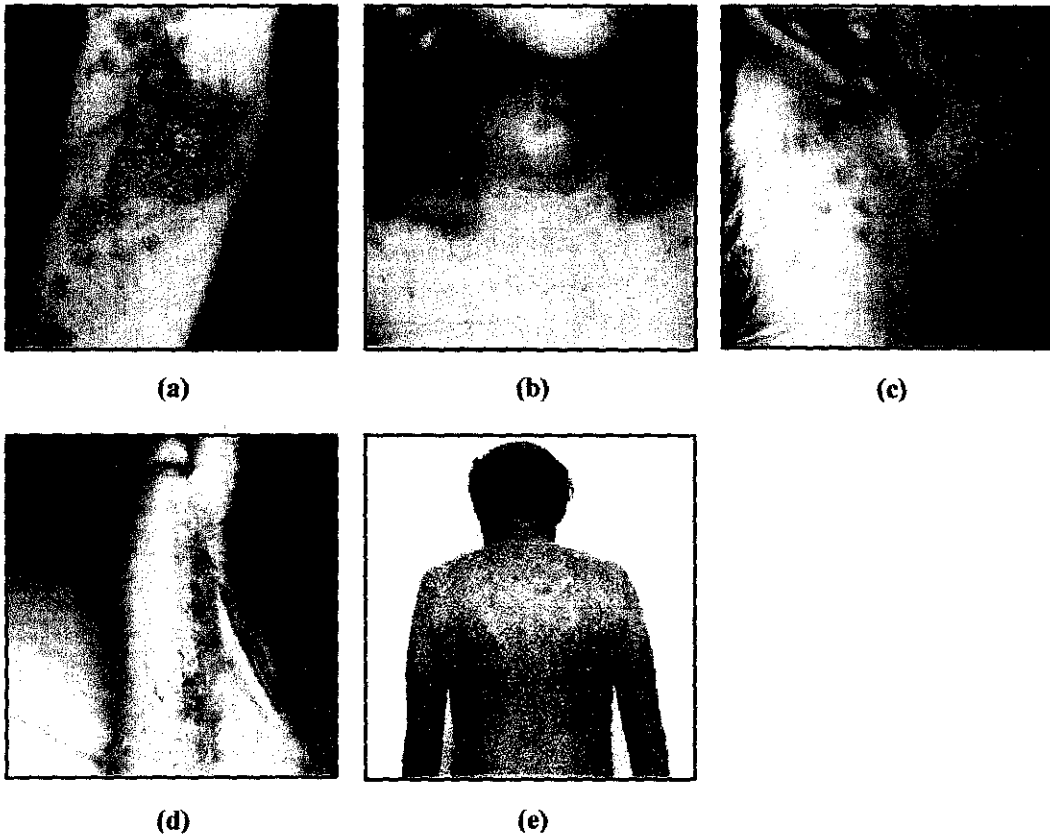
### **1.1.2 Classification of Psoriasis (Signs and Symptoms)**

The five types of psoriasis are as follows:

1. Plague Psoriasis (Psoriasis Vulgaris, Vulgar Psoriasis, Psoriasis Discoidea, Discoid Psoriasis)
  - This is the most common form of psoriasis characterized by circular- to oval-shaped red skin lesions with silvery white and scaly patches. These patches (plagues) can remain isolated, separate or join together to form a massive plague area.
2. Guttate Psoriasis (Eruptive Psoriasis, Teardrop Psoriasis, Raindrop Psoriasis)
  - It is usually triggered by a bacterial infection such as strep throat. This type of psoriasis is characterized by small, water-drop-shaped-red-spots on the skin on the trunk, arms, legs and scalp.
3. Inverse Psoriasis (Flexural Psoriasis, Psoriasis Inversa, Inverted Psoriasis, Skin Fold Psoriasis, Genital Psoriasis, Intertriginous Psoriasis, Inguinal Psoriasis and Atypical Psoriasis)
  - This type is characterized by smooth and shiny patches of red, inflamed skin in skin folds as found under the breasts and in groins and armpits. It can be most found in obese people and can be further aggravated by sweating and friction.
4. Pustular Psoriasis (Psoriasis Pustulosa, Von Zumbusch Pustular Psoriasis, Acute Pustular Psoriasis and Generalized Pustular Psoriasis)
  - This is a rare type of psoriasis characterized by raised bumps filled with non-infectious pus. The skin under and around these bumps is reddish and tender.

5. **Erythrodermic Psoriasis** (Psoriasis Universalis, Exfoliative Psoriasis, Psoriatic Erythroderma, and Psoriatic Exfoliative Erythroderma)

- This psoriasis type is characterized by red, inflamed and scaly rashes on the entire body. This skin ailment is often itchy and extremely painful.



**Figure 1: Different Types of Psoriasis**

**(a) Plaque Psoriasis, (b) Guttate Psoriasis, (c) Inverse Psoriasis, (d) Pustular Psoriasis, (e) Erythrodermic Psoriasis**

(Source: Mayo Clinic - <http://www.mayoclinic.com/health/psoriasis-pictures>)

### 1.1.3 Incidences of Psoriasis

An epidemiology study of psoriasis by the dermatology department of Hospital Tengku Ampuan Rahimah, Klang from January 2003 to December 2005 was done on 5607 outpatients [3]. The diagnosis was further confirmed following a series of physical examination and laboratory investigation during the three-year period. The findings are tabled as follows:

*Table 1: Distribution of psoriasis cases according to age and gender among outpatients attending dermatological Clinic in Hospital Tengku Ampuan Rahimah Klang, Selangor, Malaysia [3]*

Age Groups	Number of Psoriasis Patients (%)		Total Number (%)
	Male	Female	
0 - 9	16/413 (3.9)	16/491 (3.9)	32/904 (3.9)
10 - 20	38/622 (6.1)	39/628 (6.1)	77/1250 (6.2)
21 - 39	90/700 (12.9)	89/878 (12.9)	161/1578 (10.8)
40 - 60	123/540 (22.8)	73/657 (22.8)	296/1197 (17.2)
> 60	39/338 (11.5)	164/340 (4.7)	55/678 (8.1)
<b>Total</b>	<b>316/2613 (11.6)</b>	<b>215/2994 (7.2)</b>	<b>531/5607 (9.5)</b>

According to Table 1, 531 patients out of 5607 of the clinic's total number of outpatients were diagnosed with psoriasis. This equates to a staggering 9.5 % for the psoriasis incidences in that particular clinic itself.

*Table 2: Distribution of psoriasis according to ethnic groups among outpatients attending Hospital Tengku Rahimah, Klang, Selangor, Malaysia [3]*

Ethnicity	Number of Exams	Number of Psoriasis Patients (%)
Malay	2720	233 (8.6)
Chinese	737	44 (6.0)
Indian	1975	255 (12.9)
Others	175	9 (5.1)
<b>Total</b>	<b>5607</b>	<b>541 (9.6)</b>

Table 2 shows the breakdown of numbers by ethnicities and this data would come in helpful to further explain the clustering of the methodology in the later section.

*Table 3: Area of body commonly affected with psoriasis [3]*

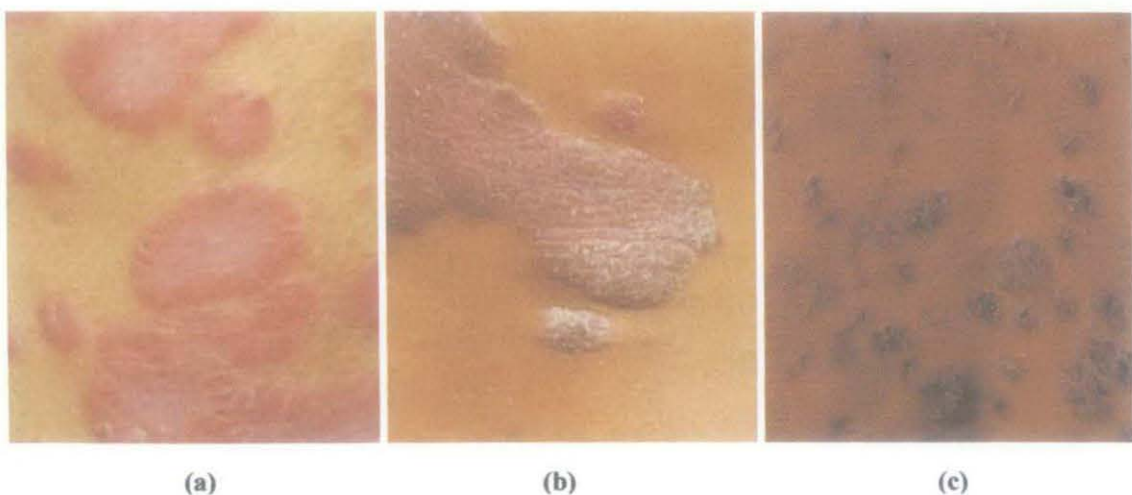
Affected Area	No. Affected	(%) Affected
Arms	287	53
Legs	238	44
Trunk	86	15.9
Genital Region	67	12.4
Scalp	64	11.8
Hands	39	7.2
Eyebrows	37	6.8
Soles	26	4.8
Face	13	2.4

These numbers are however are not indicative of the prevalence of psoriasis within Malaysia since prevalence is a different measure of the disease's occurrences. The extrapolated prevalence of psoriasis in Malaysia based on a study done by CueResearch was 475,638 (with population estimated to be 23,522,482 as obtained from US Census Bureau, International Database, 2004). [4] The statistics would yield to 2 %, which concurs with the general idea that psoriasis affects 2 - 3 % of a population.

## 1.2 Problem Statements

To compare the subjective and objective measures of reduction of psoriasis, the Department of Dermatology, Aberdeen Royal Infirmary, Foresterhill conducted a study [5]. The findings of the paper published in the year 1997 observed that subjective measures showed greater variation and relatively overestimated improvement as opposed to objectivity. By extension of these findings, assessment by clinical practitioners may not be the most accurate method for detection and assessment of psoriatic lesions.

Preliminary observations of images of lesions of random patients with *Psoriasis Vulgaris* as shown in *Figure 2* seem to show differences in colour between different lesions even for the same individual (the intra-variability of diagnostic parameters). As the skin tones differ in colour, the difference in colour of lesions (the inter-variability of diagnostic parameters) too should be taken into account. For example in fair-skin patients, the lesions manifest as bright red or pink spots but have a dark purplish or blackish hue in dark-skin patients.



**Figure 2: Different Skin Tones and Corresponding Lesion Colours**

**(a) Fair Skin, (b) Middle Skin (Brown Skin), (c) Dark Skin**

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

In view of the aforementioned problems, alternatives such as objective digital image processing algorithms have been written to assess the chromatic differences between lesions belonging to different patients and between lesions belonging to the same individual and subsequently segment the psoriatic lesions as accurately as possible. So far methods have been primarily geared toward *Psoriasis Vulgaris* lesion segmentation due to the fact that it affects 80 – 90 % of psoriasis sufferers [6] [7]. The work by Ormerod, *et al.* clearly shows that the objective digital image analysis is more accurate and has lower variations than subjective analysis measures [5]. However, the digital image analysis approach has its own set of problem, one of which is the segmentation of lesions.

Previous image analysis works using the thresholding approach have yielded considerably accurate results [8] [9] [10]. This research aims to improve on the accuracy of the thresholding works by carrying out the approach on different colour channels. Using the work of Maletti, *et al.* [11] as a telling example, the thresholding technique produced accurate results except for the dark skin tone.



(a)



(b)



(c)



(d)

**Figure 3: Original Images and Transformation into Different Colour Spaces**  
**(a) Fair Skin (Brown Skin), (b) Difference in Blue-Green Image for Fair Skin (Brown Skin),**  
**(c) Dark Skin, (d) Difference in Blue-Green Image for Dark Skin**

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

The thresholding approach was carried out in the difference in the blue-green image for fair skin, middle skin and dark skin. While the bimodality appears to be easily thresholded in the blue-green image for fair skin (as shown in *Figure 3(b)*) the same cannot be said for the dark skin image. The thresholding technique, one of the easiest approaches, yielded highly accurate results for Maletti, *et al.* in the fair skin and middle skin tone but the accuracy degenerated for the dark skin tone. The fault does not lie with the approach but rather the wrong colour channel. Other examples are further discussed in Section 2.3 under the heading of *Related Works*.

To ensure that the proposed algorithm is robust to the inter-variability and intra-variability as explained in the earlier part, three skin-and-lesion models corresponding to the three different skin tones are constructed. Each skin tone takes into account the statistical properties of all patients that fall within the particular skin tone cluster and would therefore prove to be an adequate representation to carry out the algorithm.



### 1.3 Objective

The severity of psoriasis can be assessed more accurately using the Psoriasis Area and Severity Index (PASI) [5]. The PASI scoring includes individual assessment of the area, erythema (redness), scaliness and thickness of the lesions. However, each individual assessment requires the psoriatic lesion to be first segmented from the skin sample. This report investigates digital image analysis techniques to segment psoriatic lesions.

### 1.4 Scope of Study

The scope of the research work is summarized as follows:

1. The focus of the study is based on *Psoriasis Vulgaris*, the most common type of psoriasis. Other psoriasis types are excluded.
2. The patient database for model constructing is to be obtained from 90 patients from Dermatology Clinic, Hospital Kuala Lumpur to widely model the statistical properties of various skin and lesion colours, tones and intensity
3. Skin colours would be divided into three clusters based on skin tones of patients.
4. Random distribution of the pixels in the RGB channel is used to construct the skin-and-lesion-model. Contrast adjustment would be performed if contrast between skin and lesion regions were imperceptible.
5. Average error of repeated trials conducted on each colour space would be used to determine the accuracy of each method.

## CHAPTER 2: LITERATURE REVIEW

*Psoriasis Vulgaris* affects over 80% to 90% of psoriasis sufferers. Psoriasis is a skin disorder that causes irritation and discomfort. Several image processing techniques have been developed to objectively segment the psoriatic lesions on a patient's body for improving the efficacy of treatment. Some examples include the use of a narrow bar graph partitioning method [12], variable thresholding and subimage classification [8], neuro-fuzzy approach [7] etc. In this work, an alternative approach, which combines the statistical properties obtained from a 90-patient-database followed by thresholding based segmentation technique is investigated. The discussion of the literature review would be broken down into different parts according to the flow of the methodology with references to published journals.

### 2.1 Colour Spaces

#### 2.1.1 CIE $L^* a^* b^*$

CIE  $L^* a^* b^*$  (CIELAB) colour space was designed by the International Commission on Illumination (CIE) to approximate human vision. It is perceptually uniform, denoting that the change of the same amount in a colour value would produce a change of about the same visual importance to the human eye. D.C. Tseng, *et al.* [13] conducted a study and found that the CIE  $L^* a^* b^*$  scale matches the sensitivity of human eyes with computer processing. The  $L$  component, the lightness scale closely matches the human perception of lightness. We should thus be able to segment skin tones using the  $L$  component. In Umbaugh's [14] study, he compared a number of transformations, RGB, YIQ and  $L^* a^* b^*$  to identify colours of skin tumours. He found that  $L^* a^* b^*$  transformations provided the highest diagnostic accuracy when colour data was provided to an automatic system. Dani I, *et al.* [6] used the said method for segmentation of skin and lesion areas and for elimination of unwanted excess region (hair, eye, eyebrow, nose hole, moustache, lips, beard etc. X. Yuan, *et al.* [15] compared the distribution of human skin colour in the RGB, HSI and CIELAB models and it was observed that skin colours lie inside small volume in colour space with clusters at specific points. The clusters are claimed to be more concentrated in HSI and CIE  $L^* a^* b^*$  than in RGB.

### 2.1.2 HSI

T. Carron, *et al.* [16] through his study proved that the HSI colour space was another colour space intuitive to human vision. The HSI system separates colour information of an image from its intensity transformation. Colour information is represented by hue and saturation values, while intensity, which describes the brightness of an image, is determined by the amount of light. Hue represents dominant colour as perceived by an observer. It is an attribute associated with the dominant colour as perceived by an observer. The saturation refers to the relative purity or the amount of white light mixed with a hue. The HSI colour system has a good capability of representing the colours of human perception, because human vision system can distinguish different hues easily, whereas the perception of different intensity or saturation does not imply the recognition of different colours. In *H.D.Cheng's* [17] paper, it is stated that for segmenting objects with different colours, segmentation algorithm can be applied on the hue component only by setting thresholds on the range of hues. This can be easily done on the HSI colour space but since hue, saturation and intensity values are all encoded into RGB values on the RGB colour space, it tends to be difficult. Another worthy note is that segmentation in the 1-D hue space is computationally less expensive than in the 3-D RGB space.

### 2.1.3 $I_1I_2I_3$

Ohta, *et al.* [18] analyzed more than 100 colour features obtained during segmentation of eight kinds of colour pictures. At each step of the recursive region splitting, new colour features are calculated by Karhunen-Loeve transformation of RGB. Results compared with those obtained from the Karhunen-Loeve (KL) transformations are found to be largely similar. Ohta, *et al.* found that a set of effective colour features as follows using Equation (3):

$$I_1 = (R + G + B)/3$$

$$I_2 = (R - B)/2 \tag{3}$$

$$I_3 = (2G - R - B)/4$$

An intensity histogram corresponding to the image is first obtained. Assuming that the image is composed as such that light objects and dark background (or vice versa), object and background pixels with intensity values would be grouped into two dominant modes. If the histogram obtained does not exhibit appreciable bimodality, an iterative median filter (3×3 window, 2 to 3 iterations) is applied. This set of colour was propounded through extensive systematic experiments of region segmentation. With an image, Ohta, *et al.* plotted out the weight vectors of the colour features on a plane and found out that it is possible to assume that every weight vector in the four quadrants can be approximated by the  $I_1I_2I_3$ . Comparing  $I_1I_2I_3$  with seven other colour spaces (RGB, YIQ, HSI, Nrgb, CIE XYZ, CIE L\* u\* v\* and CIE L\* a\* b\*), Ohta, *et al.* claimed that  $I_1I_2I_3$  was better in terms of quality of segmentation and the computational complexity of the transformation. (Note: transformation to other colour spaces used RGB as the base reference). Dhawan, *et al.* [19] further corroborated the stated colour features by using them for extraction of intensity and texture-based features from tumour images for segmentation. This was substantiated by Nevatia's [20] findings that results obtained by using intensity were better than RGB.

## 2.2 Global Thresholding

Sahoo, *et al.* [21] surveyed thresholding segmentation algorithms to evaluate the performance of some thresholding techniques using uniformity and shape measures. Sahoo, *et al.* categorized global thresholding into two classes: point-dependent techniques (gray level histogram based) and region-dependent techniques (modified histogram or co-occurrence based). Several different global thresholding algorithms have been developed over the years, each with its boon and bane. Uchiyama, *et al.* [22] used a sum-of-squares criterion for colour space clustering but the limitation of this method is the difficulty in representing the histogram of a colour image in a 3-D array. Ohlander, *et al.* [23] suggested a multidimensional histogram thresholding scheme using threshold values obtained from three different colour spaces (RGB, YIQ and HSI). The method by Ohlander, *et al.* used a mask for region splitting and an iterative process is executed to select the best peak of the histogram for determination of the threshold. Celenk [24] utilized the Fisher Linear Discriminant to find a line to separate the projected points for 1-D thresholding. The method operates in the CIE L\* a\* b\* colour space. The popularity of the CIE L\* a\* b\* colour space utilization is

further shown through the work by Dani, *et al.* where the method of calculating the Euclidean distance of non-involved and involved regions in the CIE  $L^* a^* b^*$  is essentially an alternative method of thresholding.

Rosenfield, *et al.* [25] used an iterative median filter to reduce the effect of noise and to improve the visibility of the borders. A  $3 \times 3$  window was used for median filtering, and 2 to 3 iterations were found to be very effective in generating filtered images suitable for segmentation. Here, histogram and an approximate colour segmentation strategy are used to identify two small windows of size  $20 \times 20$ , one inside and one outside the involved area. These windows were used later in obtaining colour variances and a proper threshold for involved and non-involved region colours. The pre-processing step was followed by a colour transformation to improve the bimodality of the image histogram.

Otsu [26] presented a nonparametric and unsupervised method of automatic threshold selection for picture segmentation by selecting thresholds through discriminant criterion, as to maximize the separability of the resultant classes. This method will be applied in this research using different colour channels to investigate its viability of skin-lesion segmentation.

### 2.3 Related Works

Jailani, *et al.* [9] captured RGB images of 81 patients using a FinePix 6900 Zoom Fujifilm digital camera which were later subjected to pre-processing. Jailani, *et al.* went with the assumption that the human skin surface was not flat and images would have undistributed intensities. Thus histogram equalization and intensity adjustment was used with global thresholding followed by morphological processing to clear up noise. One of the strong points of the work was that the images were pre-processed. In the paper, the comparison of thresholding images that were pre-processed and images that were not pre-processed were drawn out, in which both entailed differences in accuracy. One of the drawbacks of the work is that apart from removing noise, the morphological processing removed small lesions especially in the erosion step. Filters (Gaussian filters, linear spatial filtering etc.) should have been considered.

Maletti, *et al.* [11] used the difference in blue-green image to carry out the Wang's Expectation-Maximization Algorithm which provided the parameters (mean,

covariance etc.) of two Gaussian distributed classes. Maletti, *et al.* went with the assumption that the pixels of the classes are assumed to be Gaussian distributed. Once the threshold was found, a fixed discrimination function for Gaussian distributed classes was used. Segmentation in the RGB space for segmentation requires no transformation and is easy to understand. However, the three components all hold colour information while the brightness is stored implicitly as the **average** of the colour components. This makes the colour space less intuitive for image segmentation based on colour. Besides, there were instances where some of the shadows were misclassified as lesions especially in the dark skin tone. The HSI space should have been looked into since hue is invariant to shadows. [27]

## CHAPTER 3: METHODOLOGY

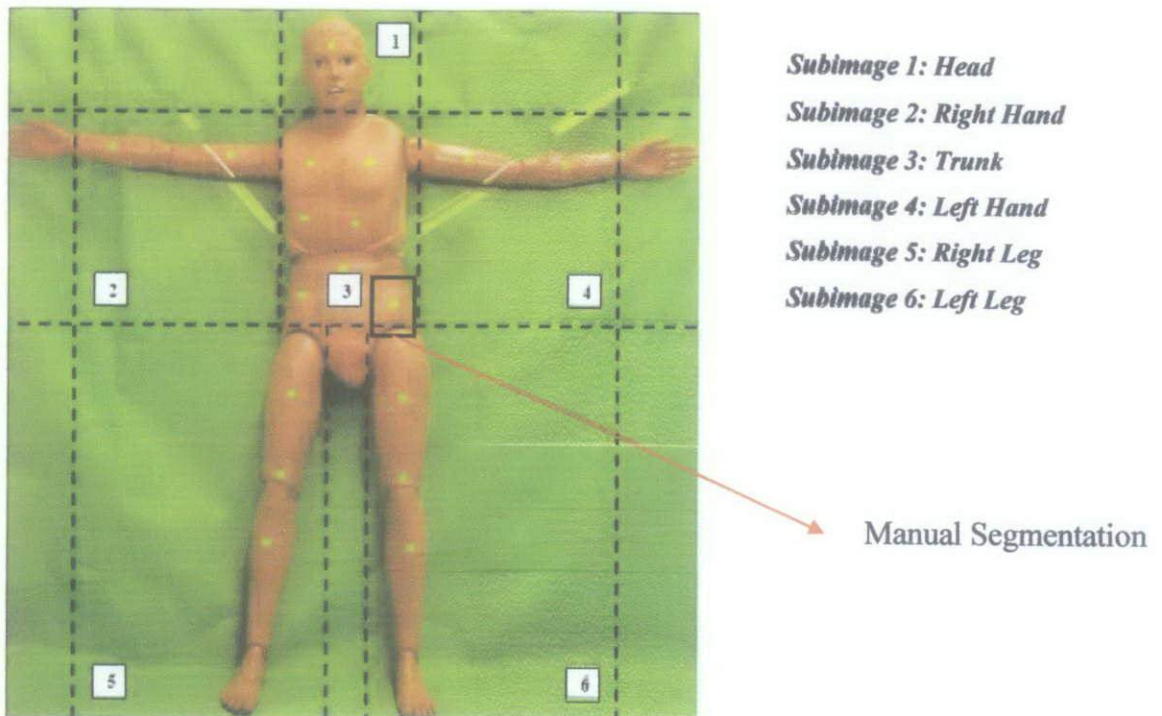
The approach taken in this research is comprised of four steps as follows:

1. Data Acquisition
2. Construction and Validation of Skin-and-Lesion-Model
3. Selection of Segmentation Techniques
4. Application of Proposed Techniques on Real Skin-Lesion Samples

The steps would be explained in detail in the subsections that follow.

### 3.1 Data Acquisition

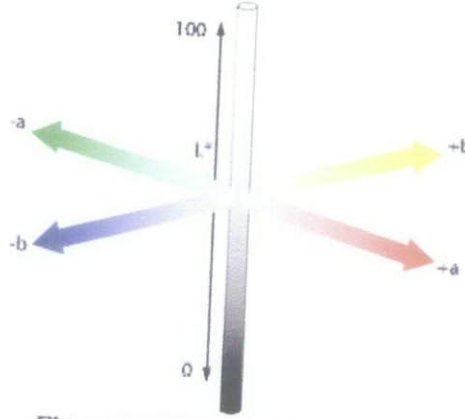
Data from 90 patients from Malay, Chinese and Indian ethnic origins with *Psoriasis Vulgaris* are obtained from the Dermatology Clinic, Hospital Kuala Lumpur. Based on the Psoriasis Area and Severity Index (PASI) which includes area, erythema (redness), scaliness and thickness assessment, the anterior posterior, right and left side of the head, trunk, right hand, left hand, trunk, right leg and left leg images were obtained using a Nikon D90 digital camera. Skin area and lesion area for each image is then manually selected using the  $\alpha$ -PASI software and stored.



**Figure 4: Segmentation on a Male Mannequin**

(Source: Centre for Intelligent Signal and Image Research, Universiti Teknologi PETRONAS)

The manually segmented images would be stored in the .xml document of the  $\alpha$ -PASI software. A MATLAB algorithm code is written to parse the .xml document and retrieve the X- and Y- coordinates of each manually-selection skin and lesion area. The skin images are stored in the sRGB format and is converted to the CIE  $L^* a^* b^*$  colour space for clustering purposes.



**Figure 5: CIE  $L^* a^* b^*$  Colour Space**

(Source: [http://dba.med.sc.edu/price/irf/Adobe\\_tg/models/cielab.html](http://dba.med.sc.edu/price/irf/Adobe_tg/models/cielab.html))

There are two steps for sRGB to CIE  $L^* a^* b^*$  transformation. Firstly, the sRGB image is transformed into CIE XYZ colour space using linear transformation given in Equation (1). The resultant CIE XYZ image is then transformed into CIE  $L^* a^* b^*$  image using Equation (2).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \times \begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} \quad (1)$$

$$L^* = 116 \left( Y/Y_n \right)^{1/3} - 16$$

$$a^* = 500 \left[ \left( X/X_n \right)^{1/3} - \left( Y/Y_n \right)^{1/3} \right] \quad (2)$$

$$b^* = 200 \left[ \left( Y/Y_n \right)^{1/3} - \left( Z/Z_n \right)^{1/3} \right]$$



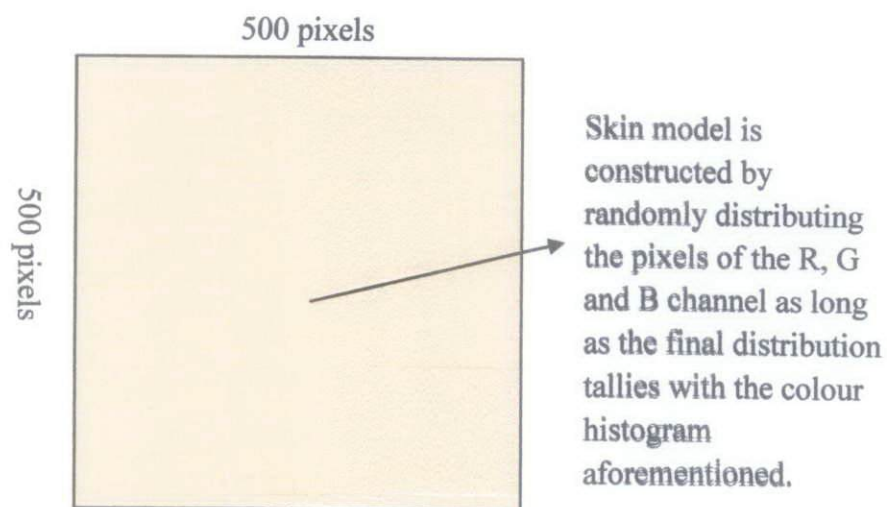
$X_n$ ,  $Y_n$  and  $Z_n$  are the tristimulus values of the reference illuminant (light source).

The parameter of interest is the lightness of the colour model which can be directly obtained through the central vertical axis  $L^*$  as opposed to the RGB colour space where lightness is dependent on relative amounts of the three colour channels.  $L^*$  values run from 0 (black) to 100 (white). The patients are separated into three clusters based on the  $L^*$  component. The reason three sectors are chosen is due to the fact of the three types of ethnicities of patients and that can be further extrapolated to cover adequately the scenario in Malaysia as shown in *Table 2* (the three major make-up percentages are from the three main ethnicities) [6].

### 3.2 Construction and Validation of Skin-and-Lesion Model

For each cluster, the R, G and B colour histograms for each image are obtained by discretization of the colours into a number of bins. The colours are then normalized considering the fact that all the images have different area values. For each channel, the number of image pixels in a particular bin for all the corresponding images in the cluster is summed up and averaged out. The same is done for all the other bins in the other images. The end result would be 3 colour histograms (each for R, G and B channel) that have taken into account the statistical properties of all the images.

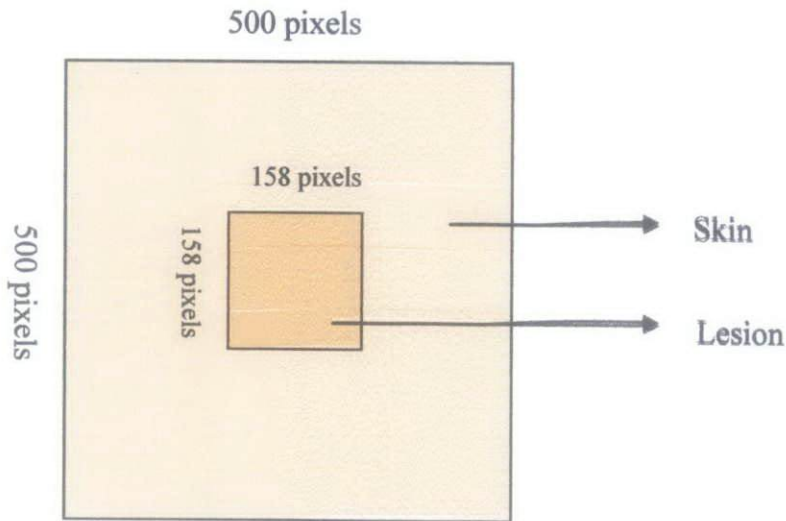
With the colour histogram, the skin model that is 500 pixels  $\times$  500 pixels is built.



*Figure 6: Skin Model*

Since the skin images have been clustered according to the lightness of  $L^*$  channel, the lesion images for each cluster chosen would thus correspond to the skin images of patients in that particular cluster as well.

The approach for the selected lesion images is the same as the binning, normalization and averaging for the selected skin images. With the colour histogram of the lesion image, the lesion model is constructed. For the initial step, lesion model is to be approximately 10 % that of the skin model, which would yield an approximate figure of  $158 \times 158$  pixels (since 10 % of skin model yields 250,000 pixels, the square root of which is 158.11 pixels).



*Figure 7: Skin-and-Lesion-Model*

**Note:** If necessary, the contrast of image can be increased by histogram equalization, adaptive histogram equalization or Contrast Limited Adaptive Histogram Equalization (CLAHE). Histogram equalization adjusts and spreads out the most frequent intensity values allowing for areas of lower contrast to gain a higher contrast. Adaptive histogram equalization computes histograms of several distinct section of the image and redistributes the lightness values of the image.

For the constructing colours of the skin-lesion models, news colours are formed through random combination of the clusters in each R, G and B component. The number of clusters in each component would be calculated using the k-means clustering approach. To minimize the overlapping of skin and lesion clusters, the fuzzy c-means method is used. This method introduces probability into the decision using the

*a priori* information (statistical nature) of the models. The results of the fuzzy c-means method are cross-referenced with the initial colour number used to construct the skin-lesion models. If the cluster of the particular colour number is the same with the fuzzy c-means results, the particular colour number (with certain R, G and B values) would remain. Otherwise it would be discarded. Each pixel for the model is checked and cross-referenced. As such, for each fair, middle and dark skin tone, the colour distribution is further cropped until two clusters are yielded. These colour numbers would be randomly assigned to construct the skin-lesion models.

For the validation, mean values of the R channel, G channel and B channel histogram of the constructed skin-and-lesion-model with the respective RGB channel histograms of the patients (for the fair-skin, middle-skin and dark-skin group) are compared. Models are validated by comparing certain parameters (mean value in this case) instead of basing it on visual perception, which is highly unreliable. Mean values are the same.

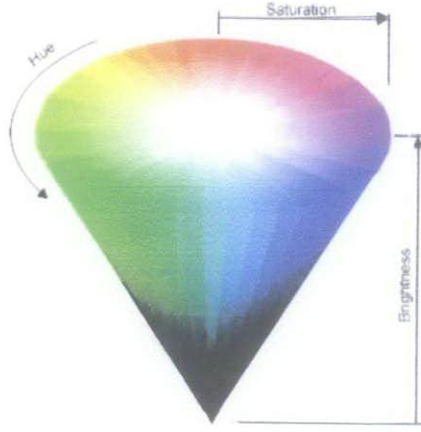
### 3.3 Selection of Segmentation Techniques

The segmentation techniques detailed in the next few sections would first be applied on the validated skin-lesion models to select the best segmentation method and colour channels before being performed on real-skin lesion samples.

#### 3.3.1 Transformation to CIE $L^* a^* b^*$ , HSI and $I_1 I_2 I_3$

Transformation from the RGB space to three different colour spaces - CIE  $L^* a^* b^*$ , HSI (hue-saturation-intensity) and  $I_1 I_2 I_3$  is applied since the hue channel is of interest. Transformations of CIE  $L^* a^* b^*$  and  $I_1 I_2 I_3$  have been presented in *Section 3.1.1* and *Section 2.1.3* respectively.

For the HSI colour space, hue is the angle between a reference line and the colour point in RGB space (ranging from  $0^\circ$  to  $360^\circ$ ), the saturation component is the radial distance from the cylinder centre and intensity is the height in the axis direction. HSI colour system is perceptually linear with human visual as the hues can be distinguished easily.



**Figure 8: HSI Colour Space**

(Source: Robotlab – [http://www.ai.rug.nl/vakinformatie/pas/index.php?content=pioneer\\_cam](http://www.ai.rug.nl/vakinformatie/pas/index.php?content=pioneer_cam))

For transformation from RGB to HSI using Equation (4),

$$H = \arctan \left( \frac{\sqrt{3} (G-B)}{(R-G) + (R-B)} \right)$$

$$Int = \frac{(R + G + B)}{3} \quad (4)$$

$$Sat = 1 - \frac{\min(R, G, B)}{I}$$

### 3.3.2 Segmentation of Lesion (Pixel-Based Thresholding)

#### *Iterative Thresholding* [28] [29]

The iterative algorithm runs along the principle as shown in the following steps:

1. Select an initial estimate for the global threshold,  $T$  based on the bimodality of the histogram.
2. Segment the image using  $T$  in

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases} ,$$

Where  $f(x, y)$  is the original image and  $g(x, y)$  is the segmented image

This will produce two groups of pixels:  $G_1$  consisting of all pixels with intensity values  $> T$ , and  $G_2$  consisting of pixels with values  $\leq T$ .

3. Compute the average (mean) intensity values  $m_1$  and  $m_2$  for the pixels in  $G_1$  and  $G_2$  respectively.
4. Compute a new threshold value:

$$T = \frac{1}{2} (m_1 + m_2)$$

5. Repeat Steps 2 through 4 until the difference between values of  $T$  in successive iterations is smaller than a predefined parameter  $\Delta T$ .

(Note: parameter  $\Delta T$  is used to control the number of iterations in situations where speed is an important issue. The initial threshold must be chosen greater than the minimum and less than maximum intensity level in the image. The average intensity of the image is a good initial choice for  $T$ .)

### ***Otsu's Method***

This method is optimum since it maximizes the between-class variance, a well-known measure used in statistical discriminant analysis. The assumption is that histogram of the image is bimodal and the image has uniform illumination such that bimodal brightness behaviour arises from object appearance only. Otsu's method has the important property that it is based entirely on computations performed on the histogram of an image, an easily obtainable 1-D array.

Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either fall in foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.

In Otsu' method the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes, is exhaustively searched:

$$\sigma^2_{\omega} = \omega_1(t)\sigma^2_1(t) + \omega_2(t)\sigma^2_2(t)$$

Weights  $\omega_i$  are the probabilities of the two classes separated by a threshold  $t$  and  $\sigma_i^2$  variances of these classes. Otsu showed that minimizing the intra-class variance is the same as maximizing inter-class variance:

$$\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t)\omega_2(t) [\mu_1(t) - \mu_2(t)]^2$$

that is expressed in terms of class probabilities  $\omega_i$  and class means  $\mu_i$  which in turn can be updated iteratively.

### 3.4 Application of Proposed Techniques on Real Skin-Lesion Samples

After *Section 3.3* the selected technique would be applied on the real skin-lesion samples. To accurately measure the accuracy of the proposed techniques, the average percentage error is used as the measurement standard. The segmented images are compared with the reference images to measure the accuracy of the proposed lesion segmentation method. Reference images are obtained by segmenting lesion area manually from the digital images as benchmark for accuracy. To calculate accuracy, the following formula is used [Fawcett, 2006]:

$$\text{Accuracy} = \frac{\text{True positive} + \text{True negative}}{\text{Total positive} + \text{Total negative}} \times 100\%$$

Sensitivity and specificity are statistical measures of the performance of lesion segmentation. Sensitivity (also called recall rate in some fields) measures the proportion of actual positives which are correctly identified as such (e.g. the area which is categorized as lesion by the suggested method and reference segmented images). Specificity measures the proportion of negatives which are correctly identified (e.g. the area which is categorized as normal skin by the suggested method and reference segmented images).

$$\text{Sensitivity} = \frac{\text{Number of true positives}}{\text{Number of true positives} + \text{Number of false negatives}}$$

$$\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of true negatives} + \text{Number of false positives}}$$

True positive: the area which is categorized as lesion by the suggested method and reference segmented images

True negative: the area which is categorized as normal skin by the suggested method and reference segmented images

False positive: the area of skin which is misclassified as lesion

False negative: the area of lesion which is misclassified as skin

Total positive: the area which is categorized by reference segmented images as psoriasis lesion

Total negative: the area which is categorized by reference segmented images as normal skin

Type I error and Type II error are also calculated. They both are precise technical terms used by statisticians to describe particular flaws in a testing process.

Type I error (error of the first kind): the wrong decision that is made when a test rejects a true null hypothesis, i.e. misclassification of lesion as skin areas

False positive rate = 1 - specificity

Type II error (error of the second kind): the wrong decision that is made when a test fails to reject a false null hypothesis, i.e. misclassification of skin as lesion areas

False negative rate = 1 - sensitivity

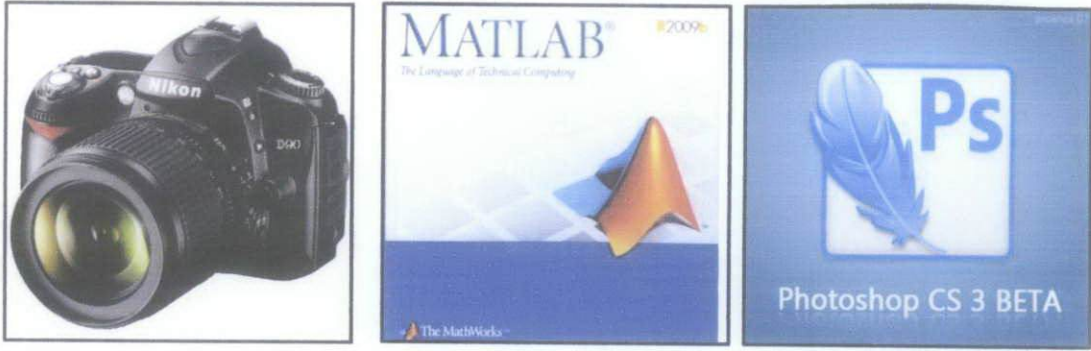
### **3.5 Manipulation of Size of Lesion**

Size of lesion is changed to 20 %, 30 %, 40 % and 50 % of skin area and resulting average error is computed.

### **3.6 Tools and Equipments Used**

The tools used in this project are a Nikon D90 camera (for capturing lesion images of patients), Adobe Photoshop (for manual segmentation of images) and MATLAB R2009b (for segmentation and image processing algorithm).



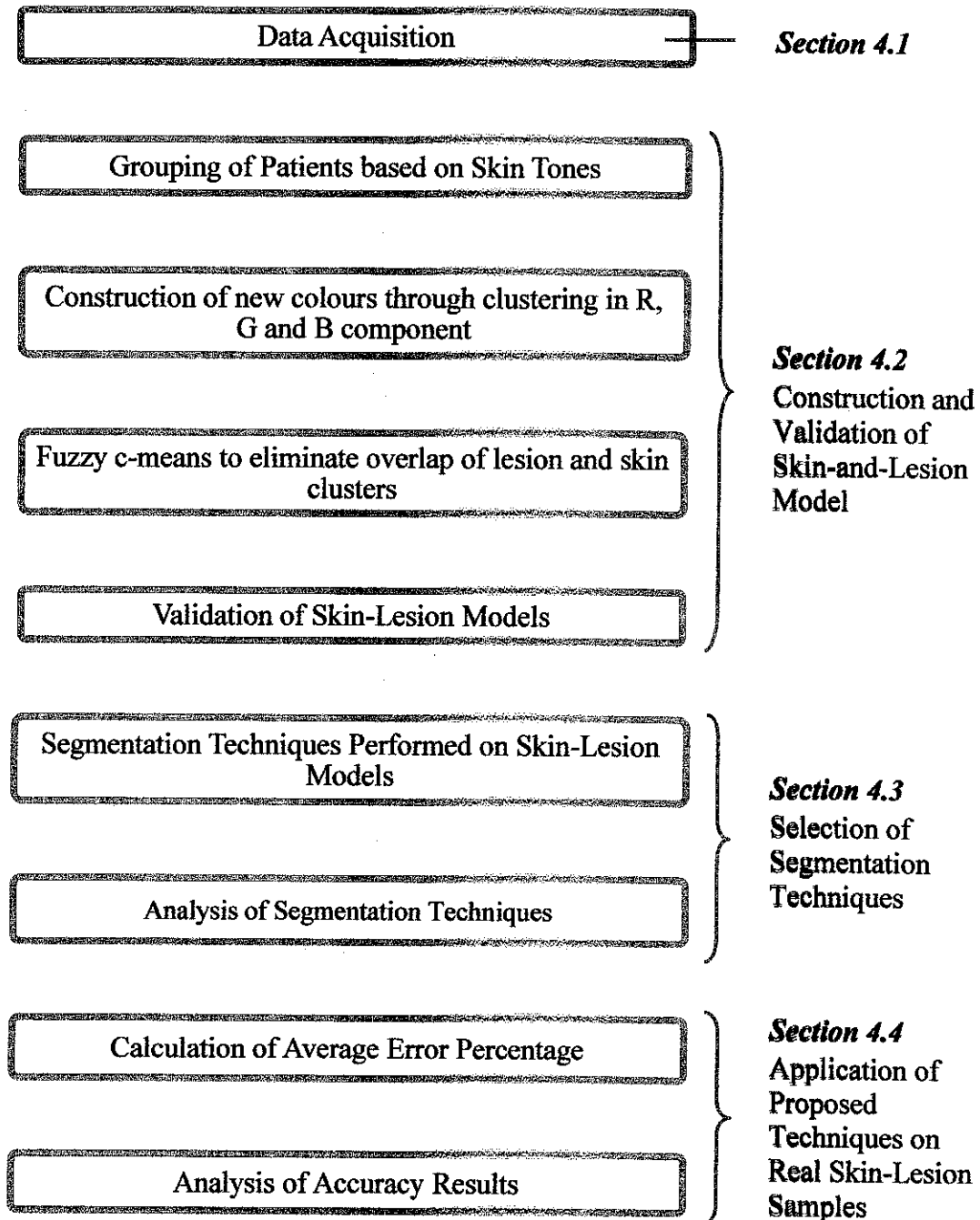


***Figure 9: Tools Used***



## CHAPTER 4: RESULTS AND DISCUSSION

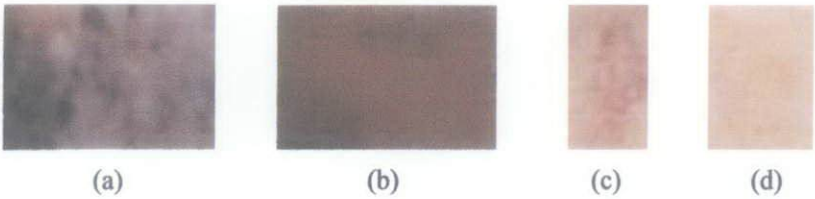
The results of this research would be broken down into sub-sections correlating to the suggested methodologies in *Section 3*. The steps are visually represented in the flow chart as follows:



*Figure 10: Flow Chart for Results and Discussion*

### 4.1 Data Acquisition

After parsing the .xml structured document, the skin and lesion samples of the 90 patients are gathered.

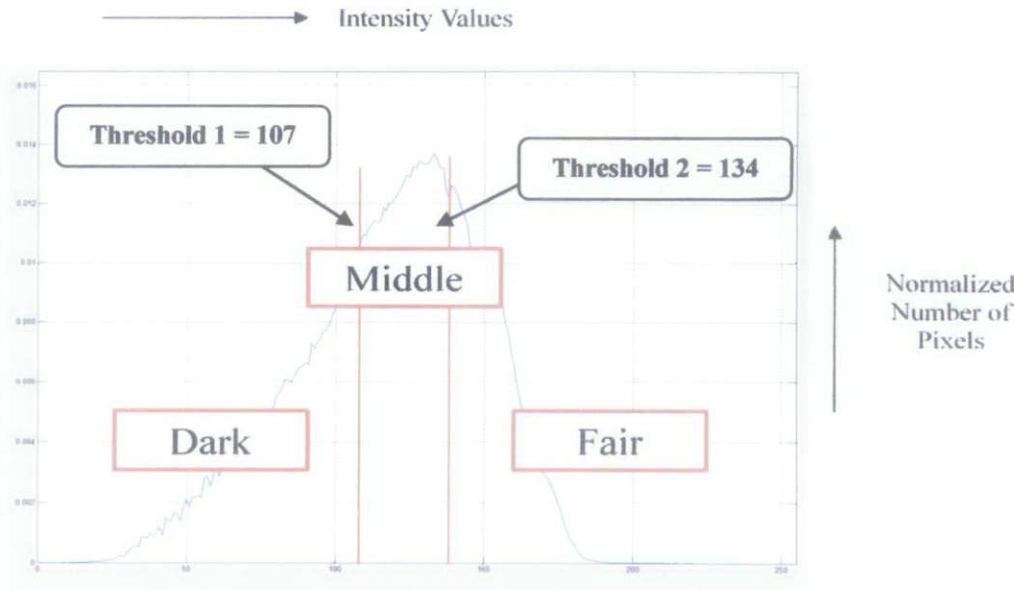


*Figure 11: Retrieval of Skin and Lesion Samples*  
(a) Lesion Sample of Patient 0008, (b) Skin Sample of Patient 0008, (c) Lesion Sample of Patient 0011, (d) Skin Sample of Patient 0011

### 4.2 Construction and Validation of Skin-and-Lesion Model

#### 4.2.1 Grouping of Patients based on Skin Tones

The values of the  $L^*$  component of the skin patches are normalized (since sizes of samples are different) resulting in one final  $L^*$  colour histogram as shown in *Figure 12*. The thresholds can be obtained through two methods: divide the total number of pixels into 3 bins of equal number of pixels or use the k-means clustering method. Both these methods produce thresholds of approximately 107 and 134 (where the red lines are drawn on *Figure 12*).



*Figure 12:  $L^*$  channel histogram for 90 patients*

Table 4: L\* Component Threshold Values

L* Component Value	Skin Tone
L < 107	Dark
107 < L < 134	Middle
L > 134	Fair

Since each of the 90 patients would have a respective mean value of the L\* component, they would be categorized them into the dark, middle or fair groups accordingly based on the range the mean values lie. Mean value of the L\* component of each patient is shown in the table below:

Table 5: Mean Value of L\* Component of Patients

Patient	Mean Value of L* Component		
1	157.8508789	29	156.9516875
2	135.2084034	30	155.3732558
3	152.8957299	31	74.14954955
4	117.1079365	32	157.6159544
5	45.71937557	33	169.3532082
6	137.509696	34	137.4666667
7	154.3851852	35	53.05489865
8	162.5923742	36	64.60157952
9	168.0433994	37	51.80077799
10	165.7615959	38	54.54207188
11	163.8942308	39	151.368932
12	164.1370391	40	128.5959596
13	167.1329787	41	146.1918605
14	159.1163019	42	142.9813953
15	149.785124	43	85.10840108
16	144.2209145	44	140.9966387
17	139.4623656	45	136.6349206
18	122.0325521	46	40.76674938
19	144.7227913	47	148.1661458
20	129.9321524	48	134.2580927
21	123.4810606	49	66.58043776
22	169.5905707	50	136.4796748
23	63.55768532	51	88.66059533
24	164.1428571	52	85.97155172
25	146.3748988	53	42.78549006
26	126.0772059	54	74.40831601
27	144.2068966	55	29.39555972
28	170.1498501	56	55.85390625
		57	94.63871817

58	69.60749118	75	57.88082647
59	143.1425432	76	57.25623322
60	133.3741365	77	56.58691207
61	136.7045855	78	153.8758065
62	141.6451737	79	80.32136491
63	148.0852632	80	154.1117958
64	145.1701389	81	136.5218703
65	160.5969511	82	168.3382353
66	144.402354	83	139.2908903
67	131.484127	84	132.4852843
68	163.8996154	85	134.0210266
69	42.1162037	86	89.34065934
70	139.6873823	87	155.4908651
71	118.0882371	88	43.49322289
72	146.7461635	89	151.7884985
73	122.4384615	90	128.1494845
74	68.78818681		

Thus, the categorization of patients results in a number of 32 patients for the fair-skin group, 30 patients for the middle-skin group and 28 patients for the dark-skin group.

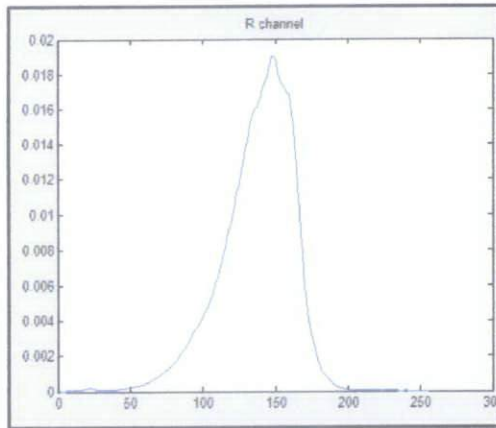
*Table 6: Categorization of 90 Patients*

Skin Tone	Patient Number
Fair Group	0004, 0006, 0011, 0012, 0013, 0015, 0016, 0017, 0018, 0019, 0020, 0027, 0029, 0030, 0034, 0035, 0036, 0038, 0039, 0045, 0047, 0053, 0069, 0070, 0071, 0074, 0078, 0084, 0086, 0088, 0093, 0095
Middle Group	0005, 0007, 0009, 0021, 0022, 0023, 0024, 0025, 0026, 0031, 0033, 0040, 0046, 0048, 0050, 0051, 0054, 0056, 0065, 0068, 0072, 0073, 0076, 0077, 0079, 0087, 0089, 0090, 0091, 0096
Dark Group	0008, 0028, 0037, 0041, 0042, 0043, 0044, 0049, 0052, 0057, 0058, 0059, 0060, 0061, 0062, 0063, 0064, 0075, 0080, 0081, 0082, 0083, 0085, 0092, 0094

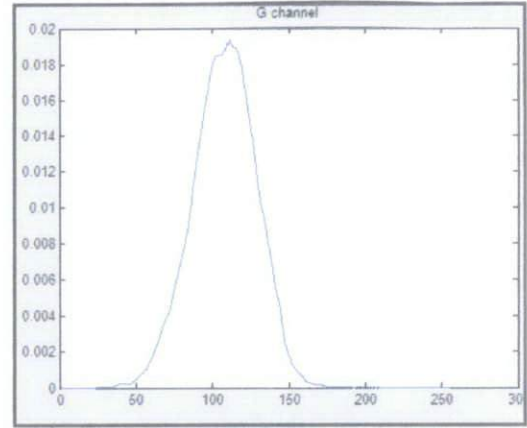
#### 4.2.2 Construction of new colours through clustering in R, G and B component

For each channel, the number of image pixels in a particular bin for all the corresponding images in the cluster is summed up and averaged out. The same is done for all the other bins in the other images. The end result would be 3 colour histograms (each for R, G and B channel) that take into account the statistical properties of all the images. Using the fair group patients as an example with a final R channel, G channel B channel histogram, the results are as shown in *Figure 13*.

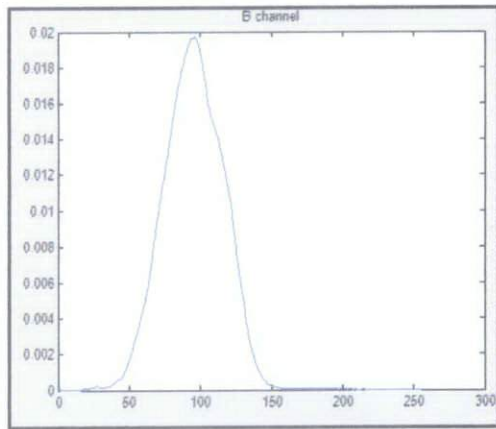




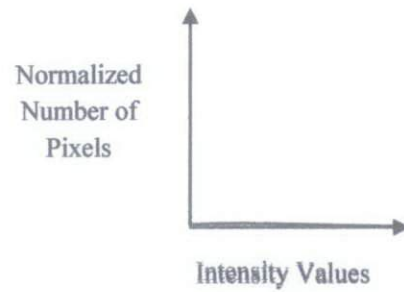
(a)



(b)



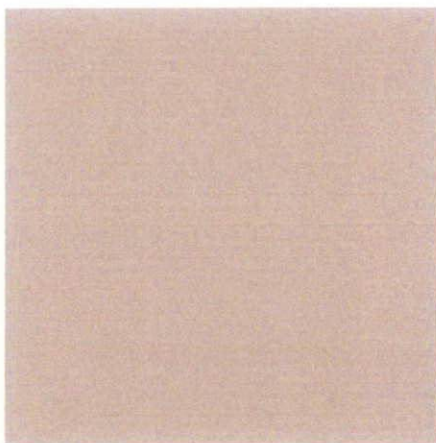
(c)



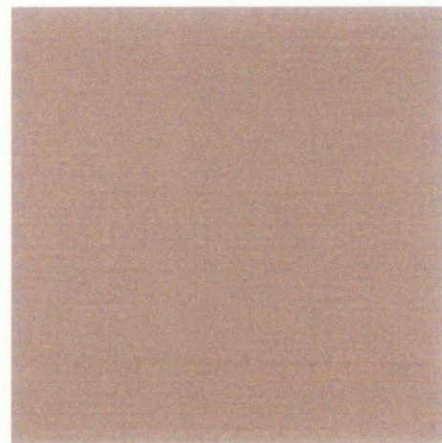
**Figure 13: RGB channel histograms for fair-skin patients**

**(a) R channel, (b) G channel, (c) B channel**

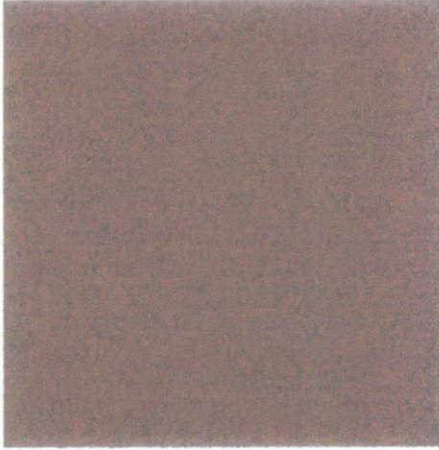
With the histograms of the colour channels,  $500 \text{ pixels} \times 500 \text{ pixels}$  skin image models are created by randomly distributing the pixels of the R, G and B channel.



(a)



(b)

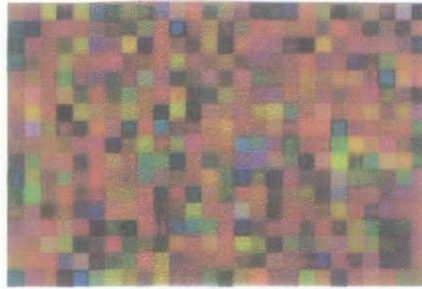


(c)

**Figure 14: Constructed Skin Model**

**(a) Fair Skin Model, (b) Middle Skin Model,  
(c) Dark Skin Model**

However when the constructed skin model (for example the dark skin model) is zoomed in,



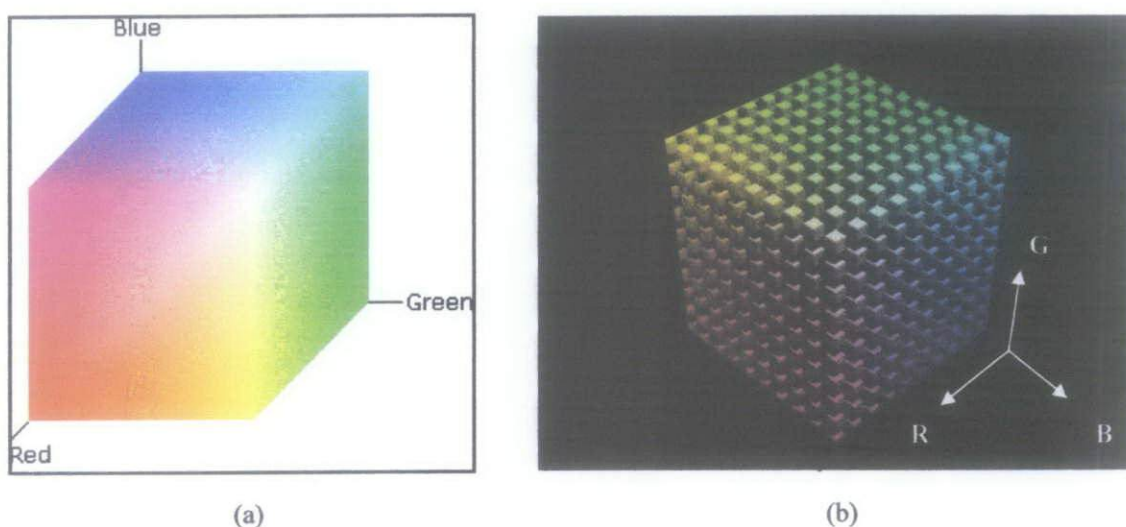
**Figure 15: Zoom-in of Constructed Skin Model**

pixels appear in several colours. This is due to the fact that random combinations of the RGB values might lead to unlikely combinations of R, G and B pixel values (unlikely denotes that the particular combination would never appear in a patch of normal skin) even though the final distribution abides by the histogram as shown in *Figure 13*. For example a pixel that has a R value of 3, G value of 123 and B value of 8 (all valid values on the RGB histogram of *Figure 13*), the resultant pixel appears as green (as shown in *Figure 16*).



**Figure 16: Pixel with  $R = 3$ ,  $G = 123$  and  $B = 8$**

Thus to minimize this, a clustering is performed on each R, G and B component to construct new colours (refer to *Figure 17(a)* for original RGB colour space). A clustering of two is performed initially, where the value of each pixel of original image for each R, G and B component is taken. If the value falls below  $255/2$  (for 2 clustering) = 122.5 for each component, the pixel is said to belong to Cluster 1 for that particular component, so on and so forth. Since each component has 2 clusters, a total of 8 new colours ( $2 \times 2 \times 2 = 8$ ) are obtained. Subsequently the random distribution of the new colours (as opposed to the RGB values) is performed on the previously stated  $500 \times 500$  pixels canvass. The number of clustering is increased until the previously stated phenomenon is not observed. Using the clustering approach, the optimum clustering per component is found to be 10 resulting in 1000 new colours (refer to *Figure 17(b)*). (This method can be verified using the k-means clustering for each colour component as well.) The R, G and B values for a particular cluster is calculated by taking the mean pixel value for that cluster in the original image.



**Figure 17: RGB Colour Spaces**

**(a) Original RGB Colour Space, (b) RGB Colour Space Segmented into 1000 Clusters**

(Source: <http://www.khyberspace.de>)

A good example to illustrate this fact is by assuming that Pixel A is of Cluster 2 in the R component, Cluster 4 in the G component and Cluster 2 in the B component. Since we have 10 clusters for each component, the range of each cluster is simply



$255/10 = 25.5$ . Cluster 2's threshold value is thus from 25.5 to 51 and Cluster 4's threshold value is 76.5 to 102. (Note that the value of 0.5 is insignificant since pixel values have to take upon whole values.). The exact value of the R component of Cluster 2 is found by calculating the mean pixel value of original image that falls within Cluster 2 of R component and assigned to Pixel A. The same is done for the G and B component cluster values. The mean values of the clusters are as shown in the tables below.

For fair-skinned patient,

*Table 7: Cluster Mean Value for Fair Normal Skin Group*

Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	87
5	NaN	112	109
6	137	133	127
7	160	150	145
8	178	168	NaN
9	196	NaN	NaN
10	217	NaN	NaN

For middle-skinned patients,

*Table 8: Cluster Mean Value for Brown Normal Skin Group*

Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	88	84
5	NaN	107	101
6	135	125	120
7	154	143	NaN
8	168	NaN	NaN
9	NaN	NaN	NaN
10	NaN	NaN	NaN



For dark-skinned patients,

*Table 9: Cluster Mean Value for Dark Normal Skin Group*

Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	32	34	32
3	54	51	50
4	76	73	70
5	100	96	92
6	121	NaN	NaN
7	144	NaN	NaN
8	NaN	NaN	NaN
9	NaN	NaN	NaN
10	NaN	NaN	NaN

*\*NaN means that no pixels fall within that cluster itself. For any new colour, that mean value would not be useful since no colour would be formed with that particular cluster in its combination*

The same method is carried out for the lesion models.

For fair-skinned patients,

*Table 10: Cluster Mean Value for Fair Lesion Group*

Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	NaN	37	38
3	58	58	57
4	79	80	81
5	103	102	102
6	127	125	126
7	153	149	149
8	176	172	173
9	196	197	196
10	226	NaN	NaN

For middle-skinned patients,

**Table 11: Cluster Mean Value for Middle Lesion Group**

Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	NaN	38	37
3	57	57	57
4	80	77	78
5	103	102	102
6	127	126	127
7	155	147	148
8	173	170	171
9	196	NaN	NaN
10	NaN	NaN	NaN

For dark-skinned patients,

**Table 12: Cluster Mean Value for Dark Lesion Group**

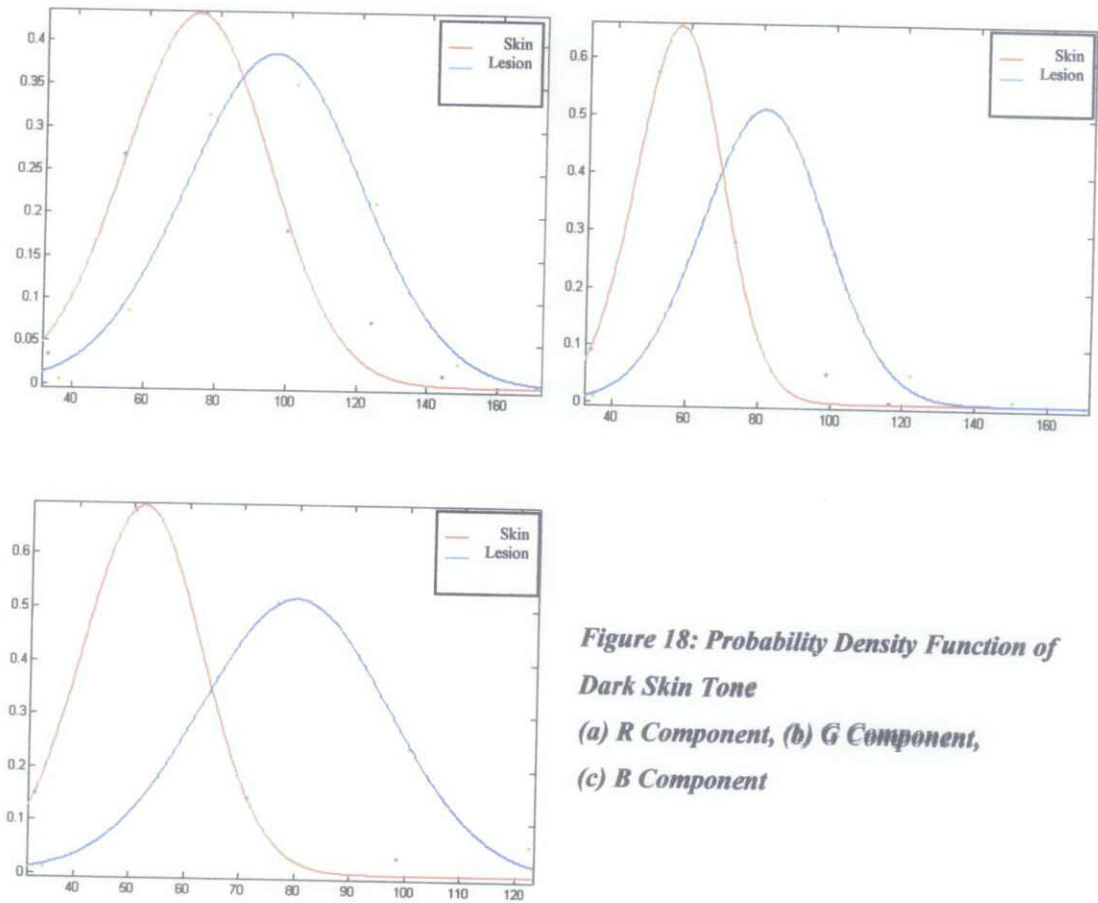
Cluster	R mean	G mean	B mean
1	NaN	NaN	NaN
2	36	35	34
3	56	56	55
4	77	77	76
5	102	100	100
6	125	122	122
7	148	150	147
8	171	170	169
9	NaN	NaN	NaN
10	NaN	NaN	NaN

*\*NaN means that no pixels fall within that cluster itself. For any new colour, that mean value would not be useful since no colour would be formed with that particular cluster in its combination.*

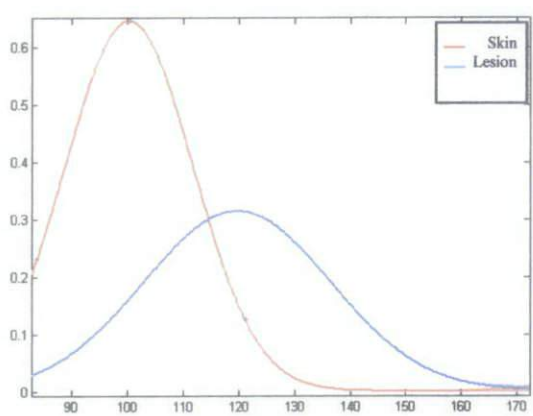
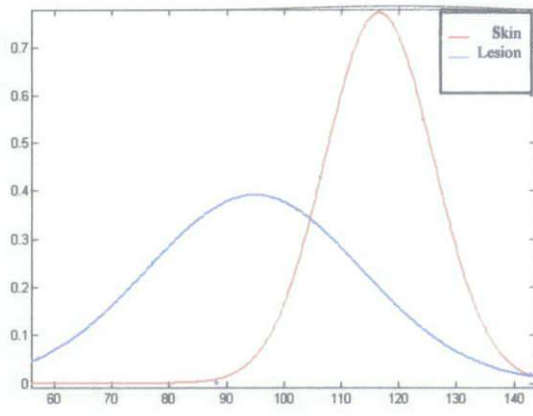
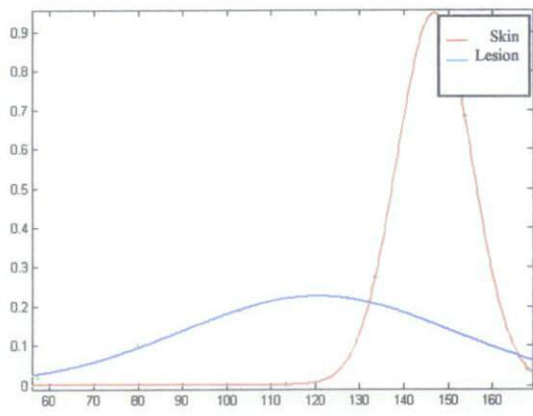
#### **4.2.3 Fuzzy c-means to eliminate overlap of lesion and skin clusters**

Once the clustering of colours has been done, 10 clusters for each RGB component would yield 1000 new colours, which would then be randomly distributed to form normal skin and lesion models for fair, brown and dark-skinned groups. The colours when plotted in 3-dimensional RGB space using MATLAB however, have overlapping of skin and lesion clusters. There are no two distinct classes of skin and

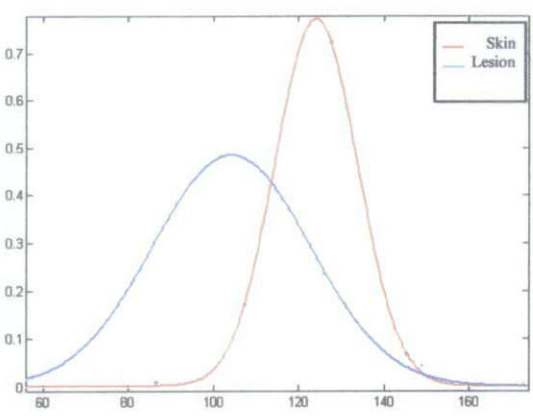
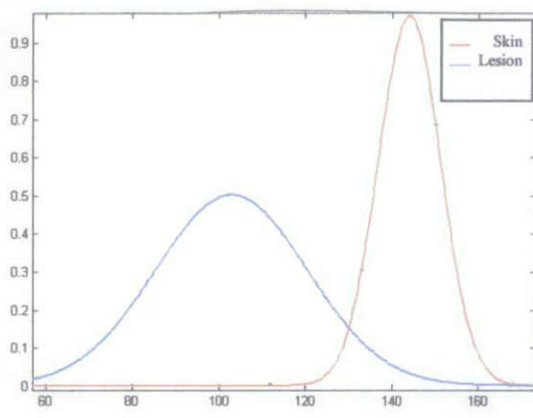
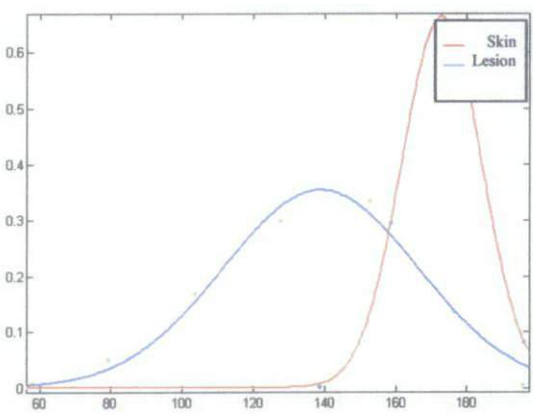
lesion colour clusters as overlaps at certain areas occur. The advantages of constructing skin-lesion models to try out the proposed algorithms are two-fold: firstly, it offers a method to accurately and adequately consolidate all statistical properties of patients within a certain skin tone and to come up with the best model representation. Secondly, it permits the model to be constantly improved to perfect the proposed algorithms before being tried out on the real skin samples. Therefore there is a need to perform slight modification on the colour numbers to take into account the overlapping of skin and lesion colour numbers and create two clusters. The fuzzy c-means method is used to overcome the problem of overlapping lesion and skin clusters. The probability density function figure for each colour component for the fair, middle and dark skin tone is shown as follows:



**Figure 18: Probability Density Function of Dark Skin Tone**  
**(a) R Component, (b) G Component,**  
**(c) B Component**



**Figure 19: Probability Density Function of Middle Skin Tone**  
**(a) R Component, (b) G Component, (c) B Component**



**Figure 20: Probability Density Function of Fair Skin Tone**  
**(a) R Component, (b) G Component, (c) B Component**

Based on the results of the fuzzy c-means probability density functions, the lesion and skin clusters can be differentiated. The colour numbers and distinct RGB values used to construct the skin-lesion model are presented in the following table.

**Table 13: Colour Number for Construction of Fair Skin-Lesion Model**

Colours	R	G	B	Type
645	152.58	101.94	101.66	Lesion
656	152.58	124.76	125.78	Lesion
756	176.12	124.76	125.78	Lesion
655	158.37	132.91	107.18	Normal Skin
766	176.76	149.94	127.50	Normal Skin

**Table 14: Colour Number for Construction of Middle Skin-Lesion Model**

Colours	R	G	B	Type
434	102.87	76.94	77.53	Lesion
534	127.14	76.94	77.53	Lesion
545	127.14	102.19	101.97	Lesion
756	172.53	126.45	126.59	Lesion
767	172.53	147.38	148.2	Lesion
544	133.38	106.21	83.63	Normal Skin
655	153.58	124.19	100.37	Normal Skin

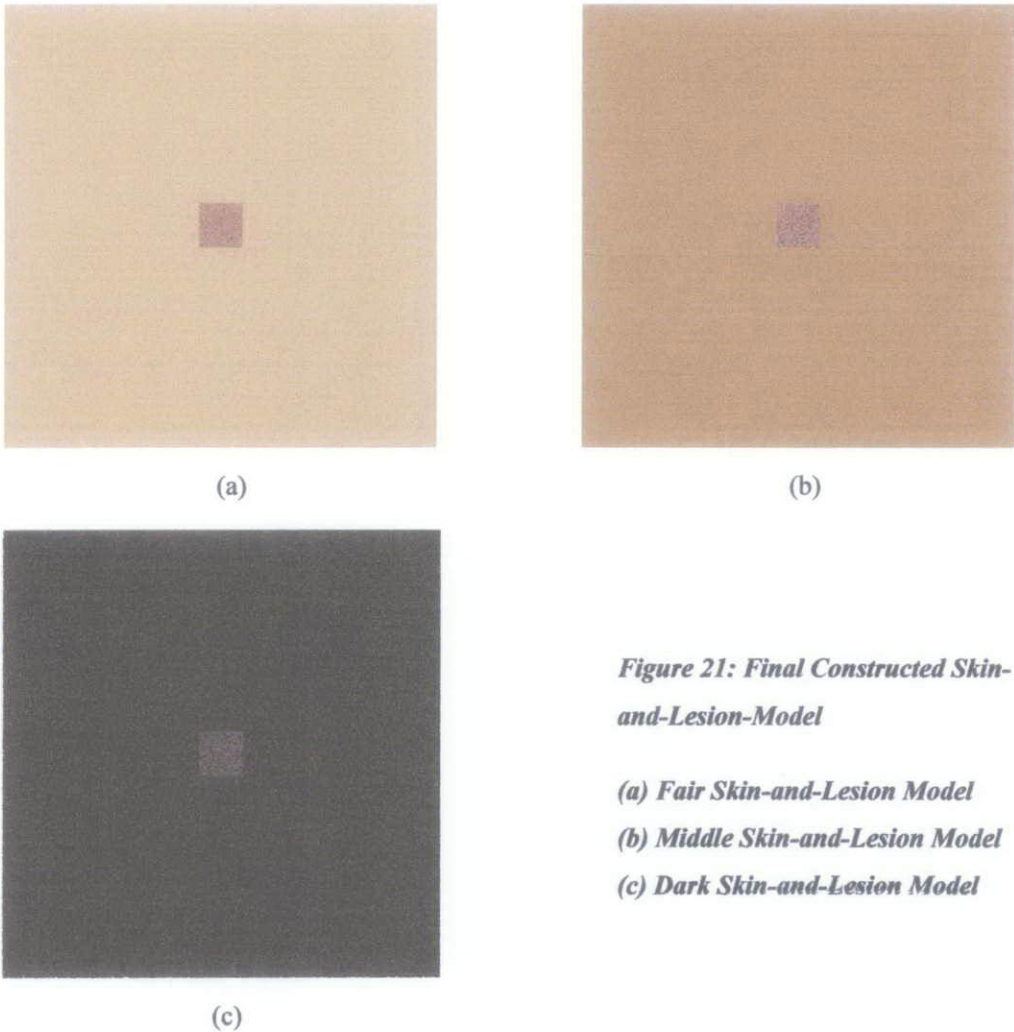
**Table 15: Colour Number for Construction of Dark Skin-Lesion Model**

Colours	R	G	B	Type
223	55.76	55.62	55.06	Lesion
323	77.29	55.62	55.06	Lesion
334	77.29	76.59	76.14	Lesion
445	102.08	100.09	100.28	Lesion
545	124.67	100.09	100.28	Lesion
433	99.8	72.85	50.06	Normal Skin



#### 4.2.4 Validation of Skin-Lesion Models

The corresponding final skin-lesion-models are as shown.



*Figure 21: Final Constructed Skin-and-Lesion-Model*

- (a) Fair Skin-and-Lesion Model*
- (b) Middle Skin-and-Lesion Model*
- (c) Dark Skin-and-Lesion Model*

For the validation, mean values of the R channel, G channel and B channel histograms of the constructed skin-and-lesion-model with the respective RGB channel histograms of the patients (for the fair-skin, middle-skin and dark-skin group) are compared. Models are validated by comparing certain parameters (mean value in this case) instead of basing it on visual perception, which is highly unreliable. The mean values of the R, G and B components of constructed skin-and-lesion models are chosen as the comparing parameters against the original skin image (instead of other colour channel components) is guided by the fact that initial construction of skin model involves clustering of colours into 10 bins within the R, G and B colour channel. It thus follows that the accuracy of the model can best be

represented by the statistical properties within the RGB colour model of each respective skin tone group.

The table below summarizes the numerical results of the mean values of skin model and corresponding original image within each skin tone group. The mean value for each R, G and B component is compared individually with that of the original image to validate the particular model. The results are as follows:

*Table 16: Validation of Skin-Lesion Model (R,G and B Mean Values)*

		Constructed Skin Model	Original Skin Image	Standard Deviation
Fair	R	170.9581	171.6402	0.482317535
	G	144.5947	143.8323	0.539098210
	B	123.8298	126.4002	1.817547270
Middle	R	148.6739	150.6137	1.371645734
	G	116.6441	117.6281	0.695793073
	B	95.943	98.0865	1.515683385
Dark	R	74.6125	77.0397	1.716289579
	G	56.9072	58.1097	0.850295904
	B	49.9998	50.0808	0.057275649

		Constructed Lesion Model	Original Lesion Image	Standard Deviation
Fair	R	141.2304	135.9621	3.725250655
	G	104.3376	102.408	1.364433245
	B	104.7468	103.578	0.826466406
Middle	R	135.7576	132.7767	2.107814604
	G	104.946	103.0898	1.312531607
	B	105.5612	103.7206	1.301500741
Dark	R	95.3892	96.3102	0.651245345
	G	79.1968	81.955	1.950341924
	B	78.4804	80.5123	1.436770269

The biggest error for either skin or lesion model, as per the highest standard deviation value between RGB mean values of original image and constructed model, is a minimal 3.72 which denotes that the model constructed is a fairly accurate representation of the original skin samples.

Each constructed skin-lesion model make-up contains a certain number of colours with distinct RGB values. Since the proposed algorithm involves thresholding

techniques performed on the CIE  $L^*a^*b^*$ ,  $I_1I_2I_3$  and HSI colour space, the RGB values can be transformed into each distinct component of the stated colour space.

*Table 17: Make-up Colours for Fair Skin Lesion Model*

Colours	$R$	$G$	$B$	$I_1$	$I_2$	$I_3$	Type
645	152.58	101.94	101.66	118.72	25.46	-12.59	Lesion
656	152.58	124.76	125.78	134.37	13.40	-7.21	Lesion
756	176.12	124.76	125.78	142.22	25.17	-13.09	Lesion
655	158.37	132.91	107.18	132.82	25.59	0.07	Skin
766	176.76	149.94	127.50	151.40	24.63	-1.09	Skin

Colours	$L$	$a$	$b$	$H$	$S$	$I$	Type
645	124	149	137	0	33	118.72	Lesion
656	141	139	132	358	18	134.37	Lesion
756	147	149	136	359	29	142.22	Lesion
655	146	135	146	31	32	132.82	Skin
766	164	136	144	28	29	151.40	Skin

*Table 18: Make-up Colours for Middle Skin Lesion Model*

Colours	$R$	$G$	$B$	$I_1$	$I_2$	$I_3$	Type
434	102.87	76.94	77.53	85.78	12.67	-6.63	Lesion
534	127.14	76.94	77.53	93.87	24.80	-12.70	Lesion
545	127.14	102.19	101.97	110.43	12.58	-6.18	Lesion
756	172.53	126.45	126.59	141.86	22.97	-11.56	Lesion
767	172.53	147.38	148.20	156.04	12.17	-6.49	Lesion
544	133.38	106.21	83.63	107.74	24.88	-1.15	Skin
655	153.58	124.19	100.37	126.04	26.60	-1.39	Skin

Colours	$L$	$a$	$b$	$H$	$S$	$I$	Type
434	91	139	132	358	25	85.78	Lesion
534	99	150	137	359	39	93.87	Lesion
545	117	138	132	0	20	110.43	Lesion
756	147	147	135	359	27	141.86	Lesion
767	162	138	131	358	15	156.04	Lesion
544	120	137	145	27	37	107.74	Skin
655	139	137	146	27	35	126.04	Skin

*Table 19: Make-up Colours for Dark Skin Lesion Model*

Colours	$R$	$G$	$B$	$I_1$	$I_2$	$I_3$	Type
223	55.76	55.62	55.06	55.48	0.35	0.11	Lesion
323	77.29	55.62	55.06	76.67	0.58	-0.07	Lesion



334	77.29	76.59	76.14	100.82	0.90	-0.54	Lesion
445	102.08	100.09	100.28	108.35	12.20	-6.19	Lesion
545	124.67	100.09	100.28	84.94	12.97	-6.26	Lesion
433	99.80	72.85	50.06	74.24	24.87	-1.04	Skin

Colours	<i>L</i>	<i>a</i>	<i>b</i>	<i>H</i>	<i>S</i>	<i>I</i>	Type
223	60	128	129	60	2	55.48	Lesion
323	66	138	132	3	29	76.67	Lesion
334	83	128	129	60	1	100.82	Lesion
445	109	129	128	0	2	108.35	Lesion
545	114	138	132	0	20	84.94	Lesion
433	86	138	146	28	50	74.24	Skin

\* Column 'Colours' represents the colour cluster number (1-1000) of which it belongs to and the value has no effect on the algorithms carried out later.

### 4.3 Selection of Segmentation Techniques

#### 4.3.1 Segmentation Techniques Performed on Skin-Lesion Models

The Otsu and Iterative Thresholding Method are carried out for the four colour spaces (RGB, CIE L\*a\*b\*, I<sub>1</sub>I<sub>2</sub>I<sub>3</sub> and HSI). Proposed algorithms are performed on the default RGB colour space to show that the RGB space is unsuitable for the two stated methods as it entails high error percentage. In the constructed skin-lesion model, the lesion has a number of 50 pixels × 50 pixels which yields 2500 pixels. Percentage error for the constructed model can thus be calculated with the formula,

$$\% Error_{lesion} = \left| \frac{2500 - Total\ Lesion\ Pixels\ Detected}{2500} \right| \times 100\%$$

$$\% Error_{skin} = \left| \frac{47500 - Total\ Skin\ Pixels\ Detected}{47500} \right| \times 100\%$$

$$Average\ \% Error_{total} = \frac{\% error_{lesion} + \% error_{skin}}{2}$$

where total lesion pixels are calculated with either Otsu or Iterative Thresholding Method using MATLAB. Total lesion pixels wrongly detected however can either be lesion pixels mistakenly assumed to be skin pixels or skin pixels mistakenly

assumed to be lesion pixels. The images of transformed skin-lesion model into the different colour channels are attached in the appendix. The table below summarizes the percentage error calculated.

*Table 20: Percentage Error by Colour Space (Fair Skin)*

	% Error Lesion		% Error Skin		Average % Error	
	Otsu	Iterative	Otsu	Iterative	Otsu	Iterative
<b>R</b>	26.53	27.32	18.86	21.69	22.70	24.51
<b>G</b>	45.03	46.70	22.86	25.47	33.95	36.09
<b>B</b>	45.03	46.70	26.75	27.87	35.89	37.29
<b>I<sub>1</sub></b>	45.03	46.70	22.86	25.65	33.95	36.18
<b>I<sub>2</sub></b>	81.50	82.30	34.90	38.23	58.20	60.27
<b>I<sub>3</sub></b>	0.80	3.40	1.20	3.40	1.00	3.40
<b>L</b>	22.28	22.90	18.86	21.32	20.57	22.11
<b>a</b>	18.50	19.63	10.86	11.23	14.68	15.43
<b>b</b>	22.68	32.80	22.72	32.80	22.70	32.80
<b>H</b>	54.97	56.30	26.56	26.56	40.77	41.43
<b>S</b>	81.50	83.43	100.00	100.00	90.75	91.72
<b>I</b>	26.53	27.00	18.86	19.25	22.70	23.13

*Table 21: Percentage Error by Colour Space (Middle Skin)*

	% Error Lesion		% Error Skin		Average % Error	
	Otsu	Iterative	Otsu	Iterative	Otsu	Iterative
<b>R</b>	16.24	18.67	100.00	100.00	58.12	59.34
<b>G</b>	77.22	79.65	46.70	48.90	61.96	64.28
<b>B</b>	59.42	59.99	77.22	77.00	68.32	68.50
<b>I<sub>1</sub></b>	77.22	79.65	46.70	48.91	61.96	64.28
<b>I<sub>2</sub></b>	36.26	56.30	29.15	31.56	32.71	43.93
<b>I<sub>3</sub></b>	0.99	3.40	1.23	3.20	1.11	3.30
<b>L</b>	77.22	79.65	46.70	50.15	61.96	64.90
<b>a</b>	41.98	45.62	36.26	39.78	39.12	42.70
<b>b</b>	1.20	2.00	1.40	2.30	1.30	2.15
<b>H</b>	36.65	39.91	29.37	32.67	33.01	36.29
<b>S</b>	24.08	34.94	21.46	24.73	22.77	29.84
<b>I</b>	77.22	79.66	46.70	48.91	61.96	64.29

*Table 22: Percentage Error by Colour Space (Dark Skin)*

	% Error Lesion		% Error Skin		Average % Error	
	Otsu	Iterative	Otsu	Iterative	Otsu	Iterative
<b>R</b>	16.08	23.67	100.00	100.00	58.04	61.84
<b>G</b>	72.98	75.45	100.00	100.00	86.49	87.73
<b>B</b>	31.74	35.37	100.00	100.00	65.87	67.69

I <sub>1</sub>	62.95	69.21	90.10	92.30	76.53	80.76
I <sub>2</sub>	0.78	3.90	1.12	3.44	0.95	3.67
I <sub>3</sub>	45.02	52.61	100.00	100.00	72.51	76.31
L	27.02	30.10	100.00	100.00	63.51	65.05
a	53.45	59.62	88.54	91.23	71.00	75.43
b	2.33	4.38	0.54	4.10	1.44	4.24
H	35.62	36.62	100.00	100.00	67.81	68.31
S	6.45	9.09	8.47	11.22	7.46	10.16
I	62.95	74.26	90.10	90.89	76.53	82.58

\* Columns highlighted with  has constant % error < 5 % for the 3 different skin models

For further simplifications, Table 23 below summarizes the corresponding colour channel for each skin-lesion model that yields below 5 % error rate.

**Table 23: Colour Channel with < 5 % Error (3 Different Skin-Lesion Models)**

		Average Error (%)	
		Otsu	Iterative Thresholding
Fair	I <sub>3</sub>	1.00	3.40
	b	0.70	2.80
Middle	I <sub>3</sub>	1.11	3.30
	b	1.30	2.15
Dark	I <sub>2</sub>	0.95	3.67
	b	1.44	4.24

#### 4.3.2 Analysis of Segmentation Techniques

Analyzing the observations made from the results from Table 20 to Table 23,

1. L\* does not work for segmentation of skin and lesion primarily since that the skin and lesion patches of a particular patient are exposed to the same amount of light or same tone, which thereby yield uniform L\* values. Thus this lacks a good bimodality to perform thresholding algorithms.
2. The Otsu's method constantly yields a lower average error percentage as compared to the iterative thresholding method for all instances. In theory, the iterative thresholding should work better, if not as well as the Otsu's method since iterations is performed until the optimum threshold is located.



At iteration  $n$ , a new threshold  $T_n$  is established using the foreground and background means:

$$T_{opt} = \lim_{n \rightarrow \infty} T_n$$

$$T_{n+1} = \frac{T_n (\text{previous}) + T_n (\text{present})}{2}$$

However through the algorithm performed on the skin-lesion models, it is found that the iteration is terminated when the change becomes small. This observation tallies with the studies of Riddler [28] and Majid, *et al.* [29] which state that *'In practice, however, iterations terminate when the change  $|T_n - T_{n+1}|$  becomes sufficiently small.'* Otsu's method on the other hand, minimizes the weighted sum of within-class variances of the foreground and background pixels to establish an optimum threshold, and thus maximizes the between-class scatter. [30]

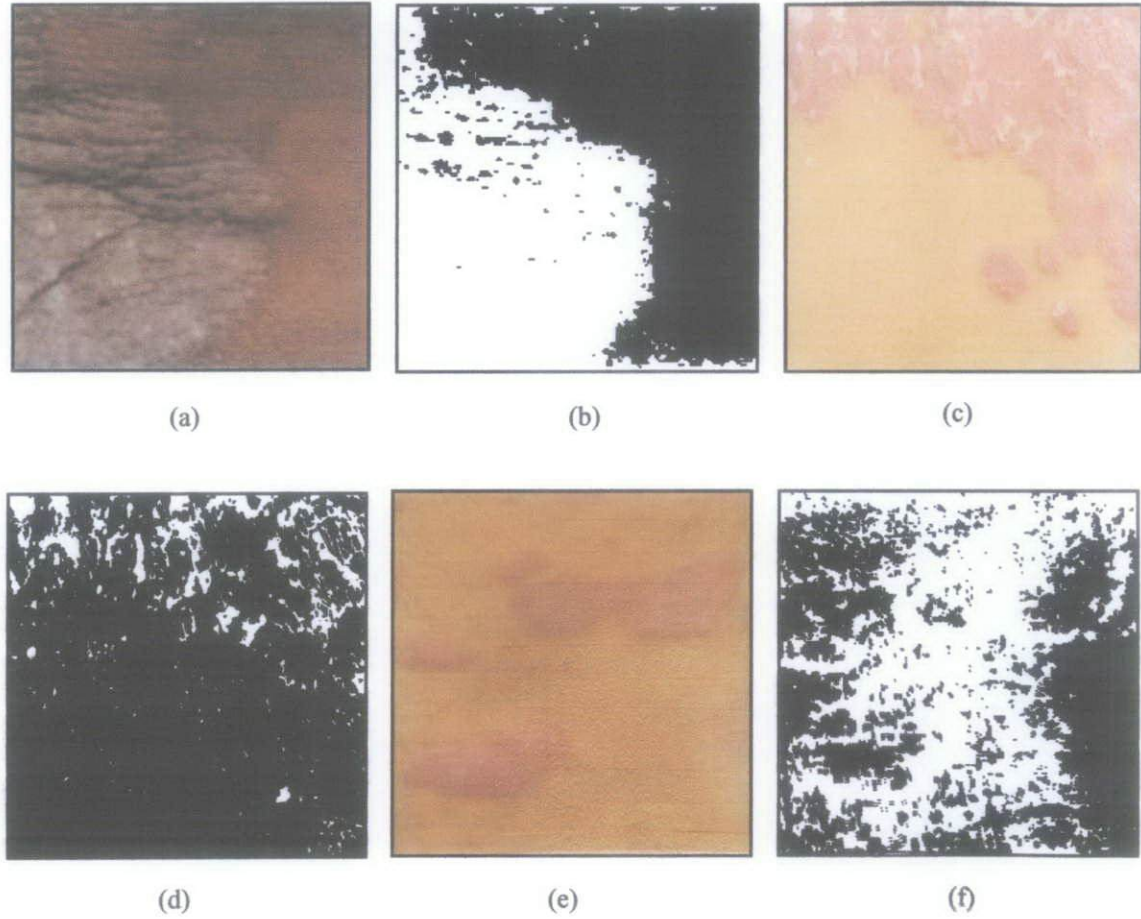
3. The global thresholding method does not work for the HSI colour space with error percentages greater than 5 %. If the error percentages between the  $H$ ,  $S$  or  $I$  colour channel are compared,  $S$  colour channel provides the lowest error percentage error for the fair, middle and dark skin-lesion models, with

fair skin-lesion model	=	90.75 %
middle skin-lesion model	=	22.77 %
dark skin-lesion model	=	7.46 %

\* Note: Error percentages shown are percentages for thresholding using the Otsu's method. Iterative thresholding yields higher error percentages – refer to explanation 2 above.

Hue and intensity are shown to be not ideal channels to carry out the thresholding segmentation algorithm for skin lesions. The suggested approach in the beginning of this report was to perform thresholding in the saturation differences in skin and lesion patches. However, as per the average percentage error of the three different skin-lesion models, the results show that only the dark skin-lesion model can be thresholded effectively

through the differences in saturation. There lacks a distinct bimodality for the fair and middle skin-lesion models. Closer inspection of the fair and middle skin-lesion models constructed reveals that the saturation values are slightly lower for the lesion, but the inter-class variance between lesion and skin is minimal. When the intensity drops, the saturation values decreases as well (shown in *Table 17*, *Table 18* and *Table 19*). In order to demonstrate the effectiveness of the Otsu's method on the fair, middle and dark skin patches, a skin and lesion image of a particular patient is thresholded using the differences in saturation.



**Figure 22: Thresholding through Differences of Saturation for Skin-Lesion Images**

**(a) Dark Skin Lesion 01, (b) Dark Skin Lesion 01 Thresholded (Otsu), (c) Fair Skin Lesion, (d) Fair Skin Lesion Thresholded (Otsu), (e) Middle Skin Lesion, (f) Middle Skin Lesion Thresholded (Otsu)**

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)



As from *Figure 22*, it is shown that through visual perception, only the dark skin model can be thresholded effectively if the differences in saturation are taken into account. The images in *Figure 22* are only thresholded with Otsu's method as it has been proven to yield lower percentage errors than the iterative thresholding method, which has been explained on *page 40*. The images only serve in this section for visual comparison of differences. The accurate percentage of error (in terms of pixels) would be looked into in the next few pages. The following few pages would also shed light on how the above images are classified as fair, middle or dark.

*Step 3.5* on *page 24* states that the lesions within the constructed skin-lesion model would be increased to different sizes and resulting average error would be computed. Separate computations similar to steps shown in the previous pages were carried out for lesions of different sizes. Observations of the tabulated average errors revealed that different sizes of lesions have the same average error values. The reason behind this observation is that the collective make-up colours of the lesions remain the same regardless of the size. To clarify this point, consider *Table 17* which shows that the make-up colours of the lesions are colours 645, 656 and 756 for lesion of size 10%. When the lesion is manually constructed to be 20% in size, the make-up colours still remain the same, namely that of 645, 656 and 756. Thus the threshold values are found to be the same. The subsequent algorithms carried out for the different-sized lesions (lesion sizes of 20%, 30%, 40% and 50%) effectively give the same average error. Thus the size of the lesion becomes insignificant but the make-up colours and thresholding algorithms are of higher interest.

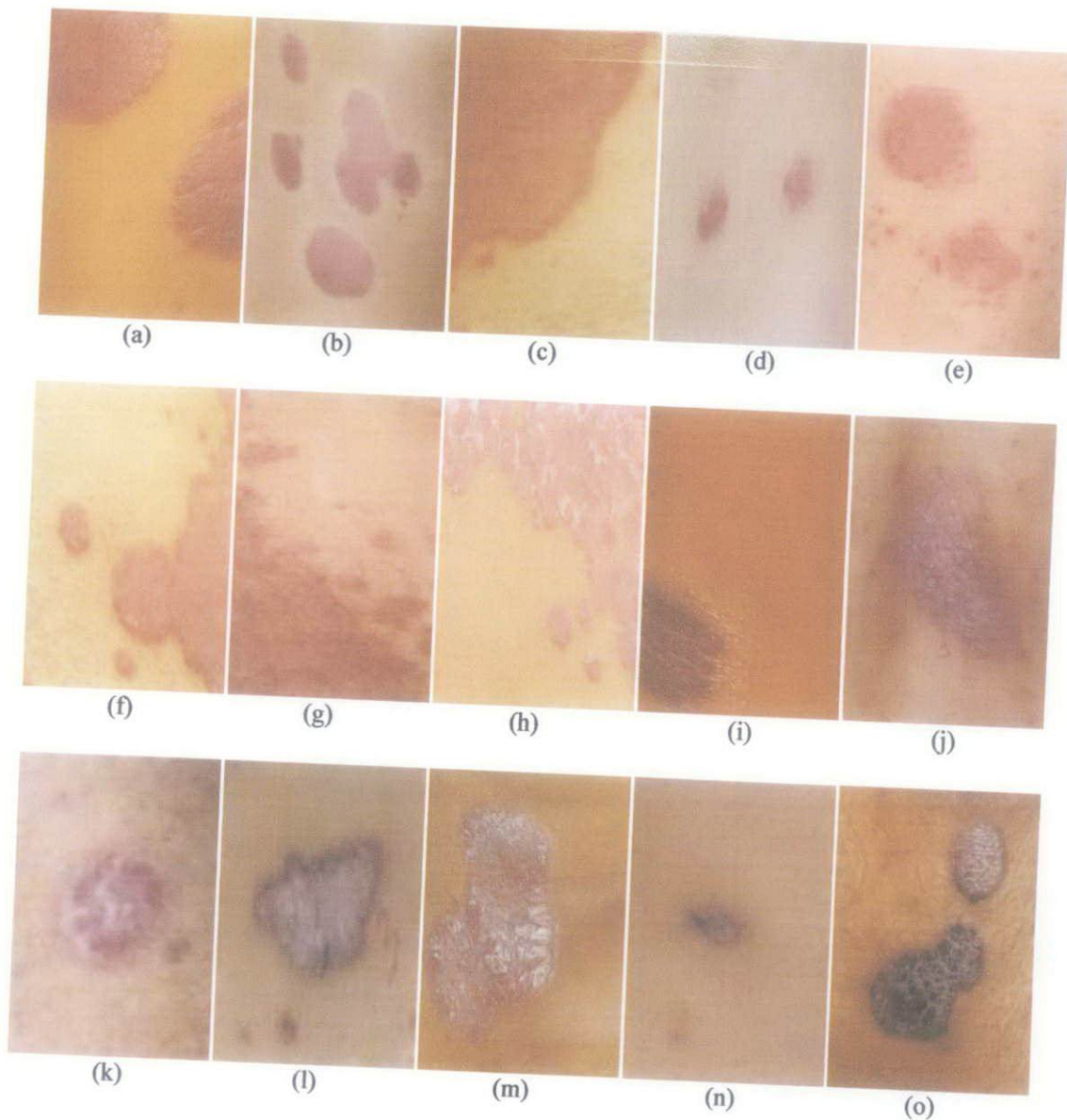
Through the construction of fair, middle and dark skin-lesion models and thorough analysis of the average error for images thresholded via two different methods in several colour channels, the report from here on in would look into the following:

**Table 24: Skin-Lesion Tones and Appropriate Colour Channels**

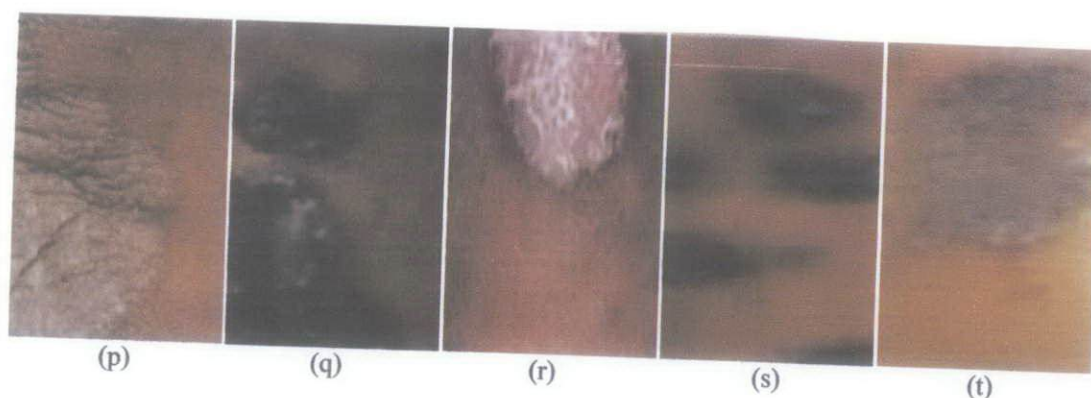
Fair Skin-Lesion	I <sub>3</sub> colour channel (Otsu's)
Middle Skin-Lesion	I <sub>3</sub> and <i>b</i> colour channels (Otsu's)
Dark Skin-Lesion	I <sub>2</sub> , <i>b</i> and <i>S</i> colour channels (Otsu's)

The Otsu's method would be carried out on the different colour channels corresponding to the skin tone stated above. The percentage of error would then be calculated by comparing the thresholded image with that of the manually segmented image. The suggested method is tried out on 20 real images of random patients sourced from Dermatology Clinic, Hospital Kuala Lumpur.

#### 4.4 Application of Proposed Techniques on Real Skin-Lesion Samples







**Figure 23: Real Skin-Lesion Patches (Patch 1 - Patch 20)**

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

The first step is to perform pre-processing on the images. Morphological processing step has shown to remove noise as well as small lesions [9]. Thus an iterative median filter, a nonlinear digital filtering technique, is used to remove noise and to improve the visibility of the borders. The aforementioned filter is a more robust average than the mean as a single very unrepresentative pixel in a neighbourhood will not affect the median value significantly. A 3 x 3 window was used for median filtering, and 2 to 3 iterations were found to be very effective in generating filtered images suitable for segmentation.

The second step is to manually segment a certain section of the skin, transform it into the CIELAB colour space and obtain the mean  $L^*$  values. The skin tones of the patches are categorized according to the thresholds shown in *Figure 12*. The mean  $L^*$  values and results of the categorization are as follows:

**Table 25: Mean  $L^*$  Value and Skin Tone for Skin Lesion Patches**

Skin-Lesion Patch #	Mean $L^*$ Value	Skin Tone
1	143.7267	Fair
2	135.4292	Fair
3	149.5061	Fair
4	155.6351	Fair
5	156.0707	Fair
6	164.5010	Fair
7	165.0517	Fair
8	172.9929	Fair
9	114.9587	Middle
10	114.3235	Middle

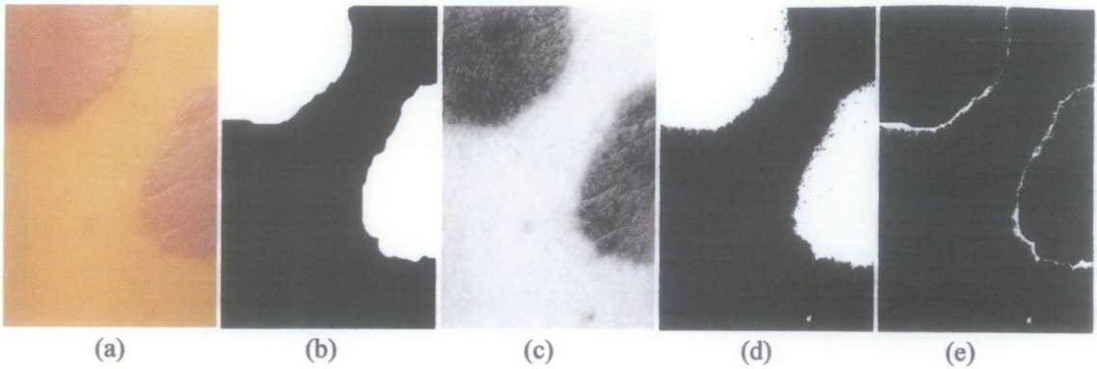


11	115.8412	Middle
12	127.2275	Middle
13	132.5018	Middle
14	133.2611	Middle
15	98.755	Dark
16	77.0229	Dark
17	70.2328	Dark
18	73.3396	Dark
19	83.2857	Dark
20	105.2402	Dark

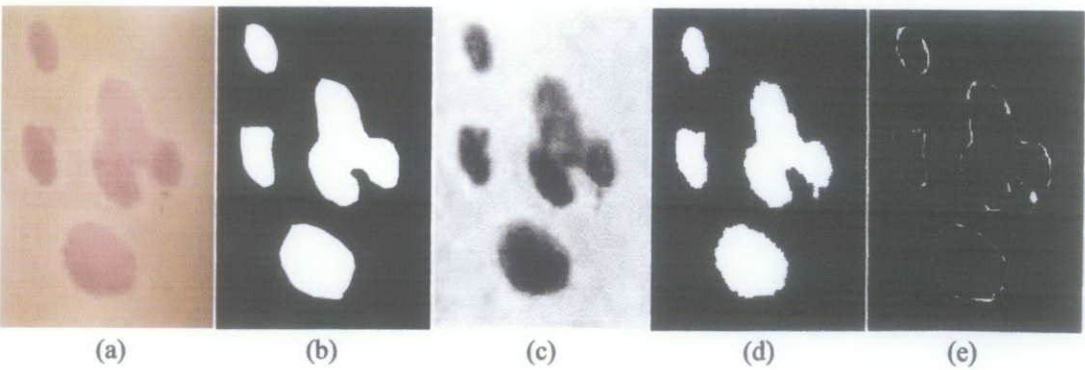
#### 4.4.1 Calculation of Average Error Percentage

The Otsu's thresholding method is carried out on the different colour channels for the skin samples corresponding to the skin tone it belongs to. The choice of colour channel for each skin tone again can be referred to *Table 24*. Image (b) of all the following figures in *Figure 24*, *Figure 25* and *Figure 26* is the image of lesion segmented manually. The percentage of difference is calculated by comparing the manually segmented images and proposed algorithms in the proposed colour channels. The percentage of accuracy is then presented in *Table 26*.

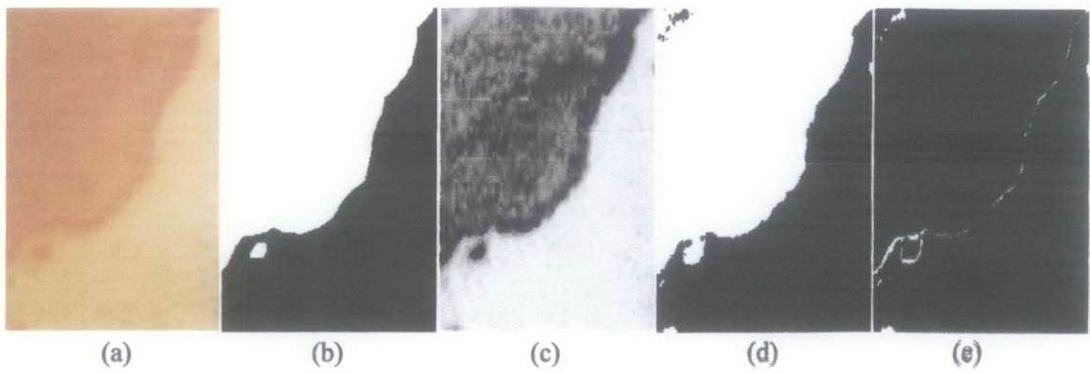
##### Patch 1



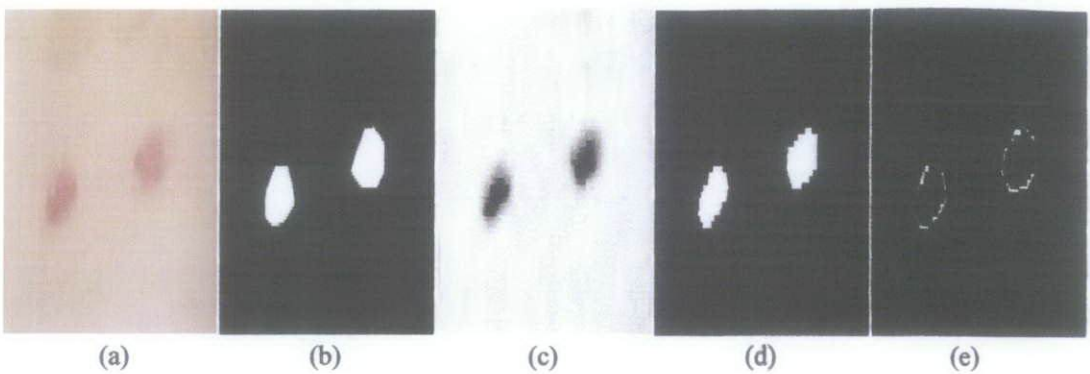
##### Patch 2



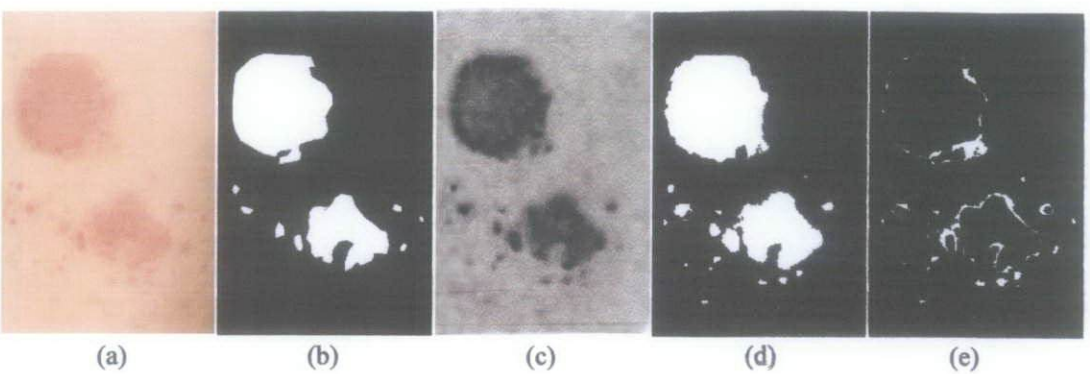
**Patch 3**



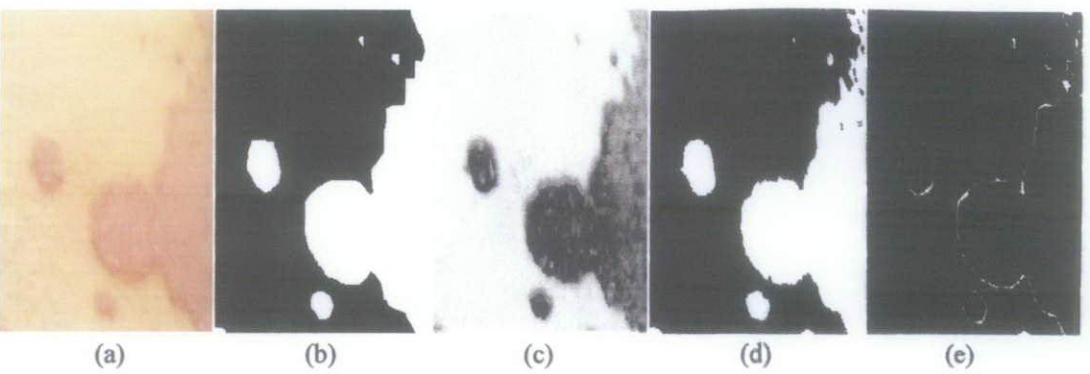
**Patch 4**



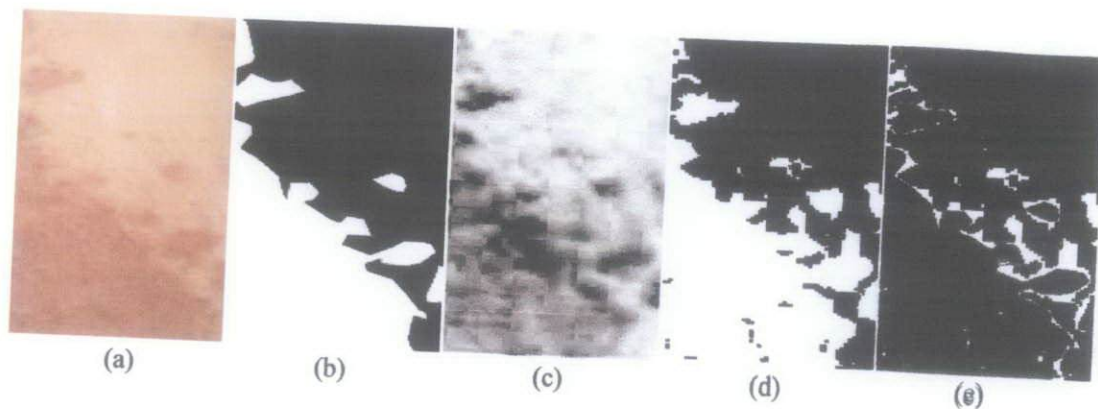
**Patch 5**



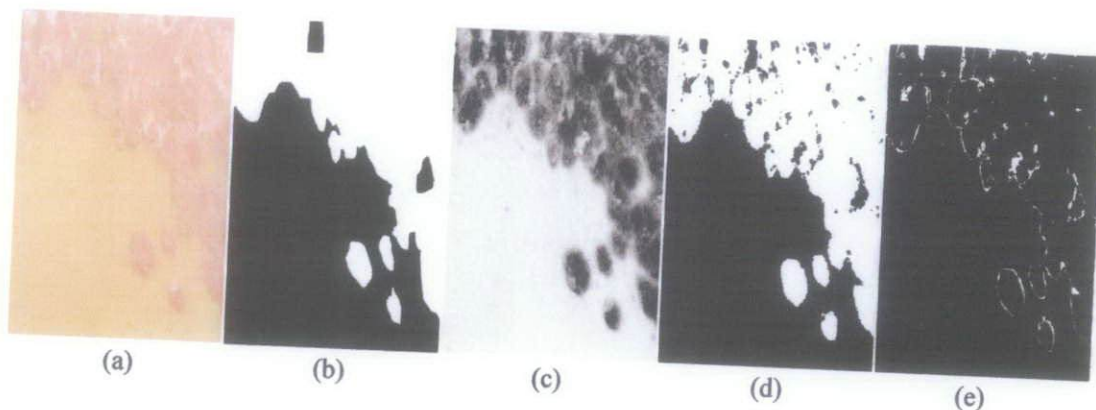
**Patch 6**



### Patch 7



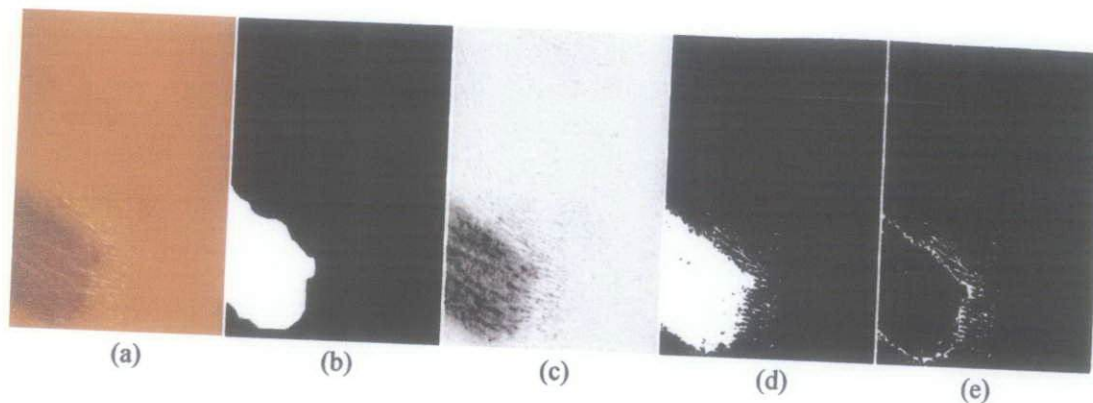
### Patch 8



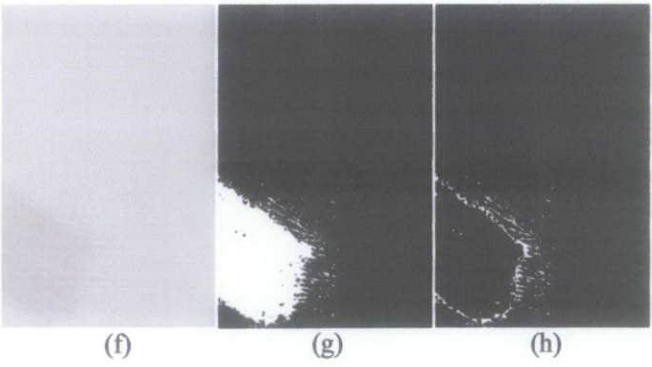
**Figure 24: Real Skin-Lesion Patches (Fair Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 1 - Patch 8)**

(a) Original Image, (b) Manual Segmentation, (c) Transformation to  $I_3$  colour space, (d) Thresholded in  $I_3$  space with Otsu's Method, (e) Differences between Manual Segmentation and  $I_3$  space

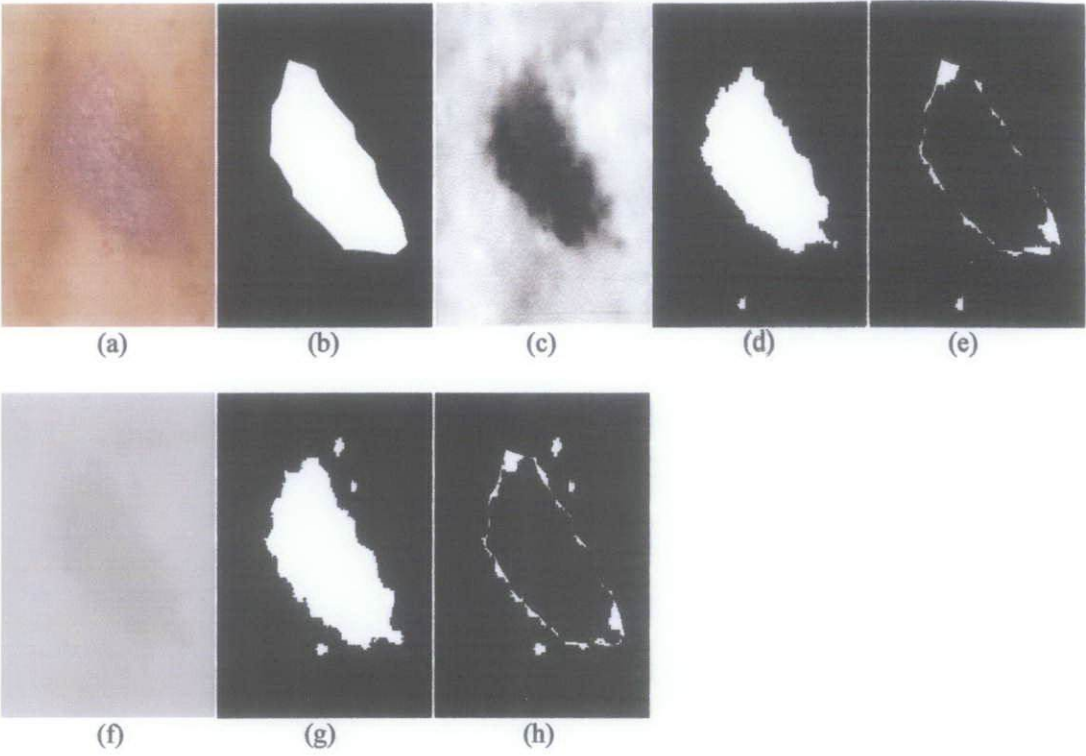
### Patch 9



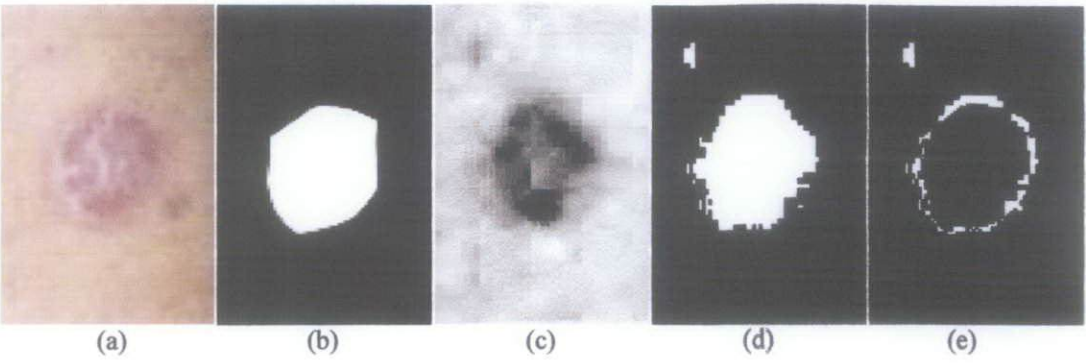


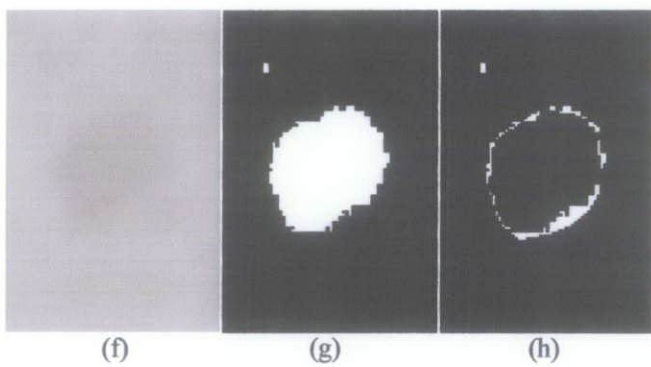


**Patch 10**

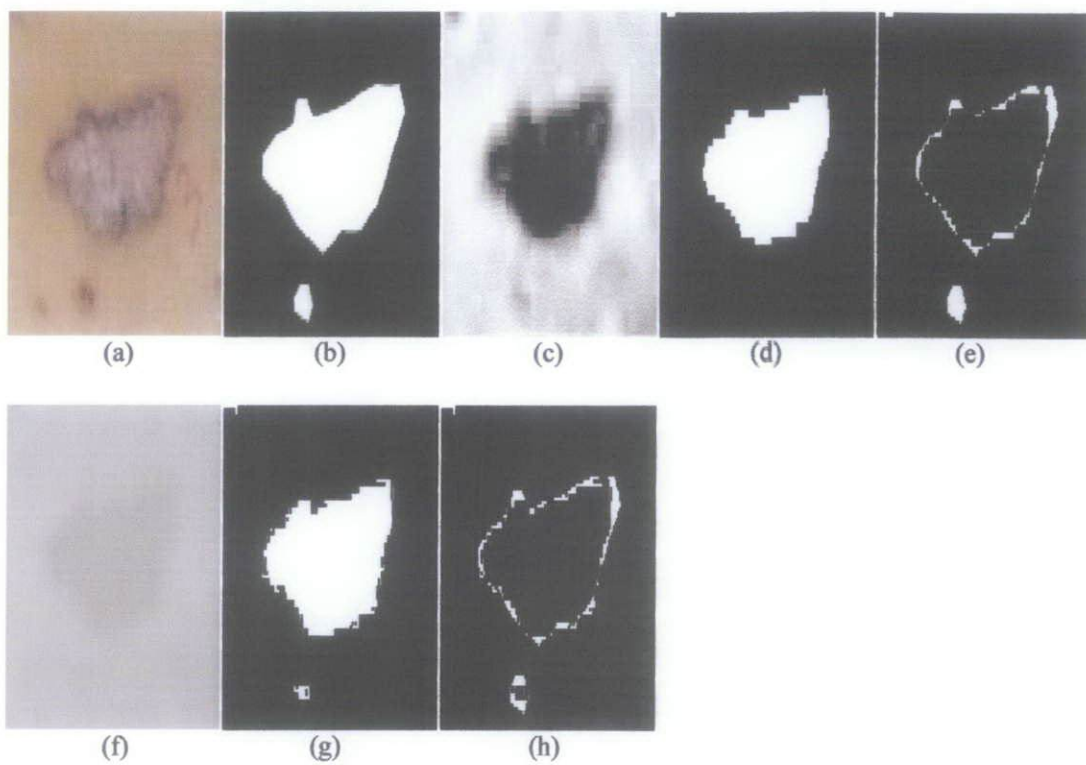


**Patch 11**

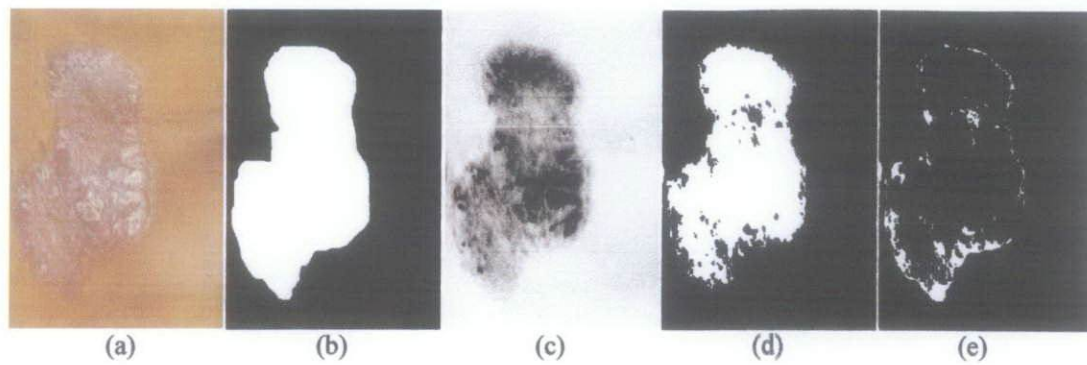


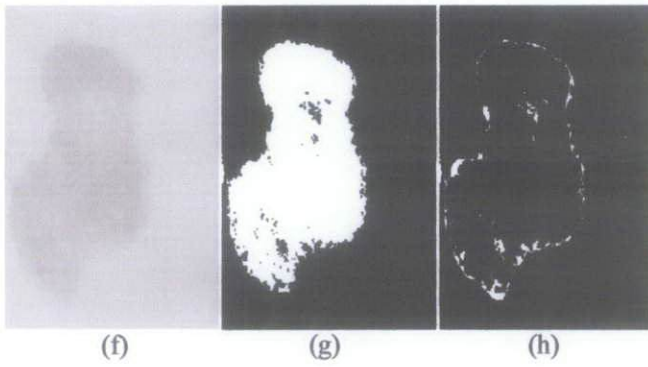


### Patch 12

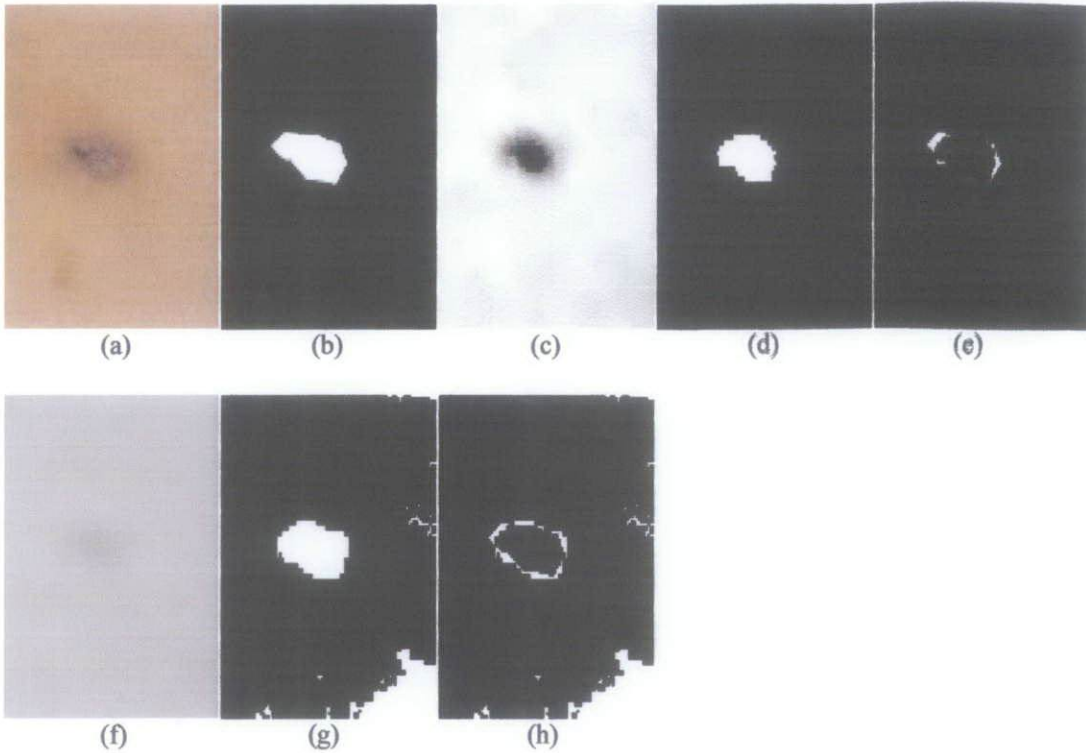


### Patch 13





#### Patch 14

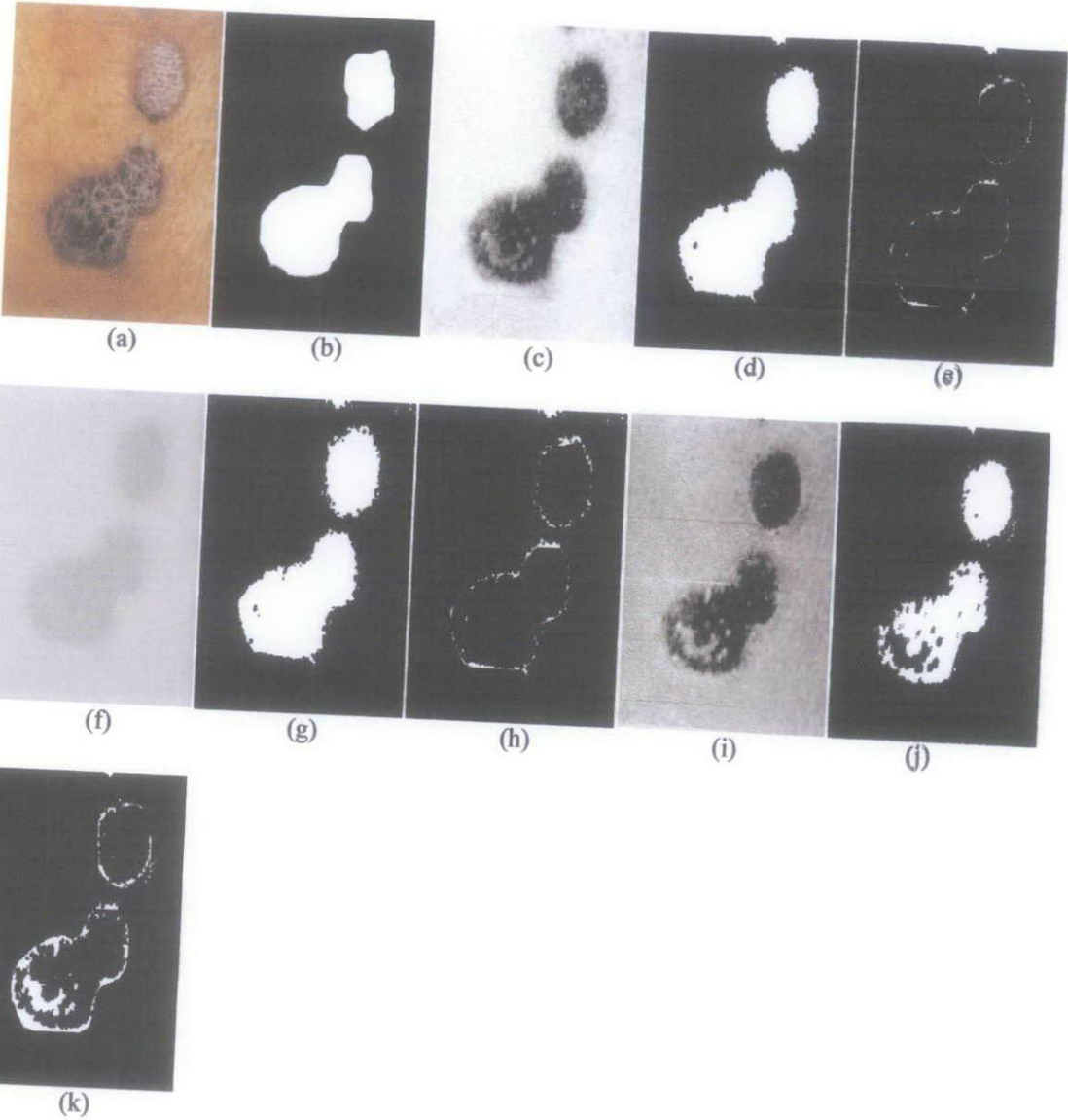


**Figure 25: Real Skin-Lesion Patches (Middle Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 9 - Patch 14)**

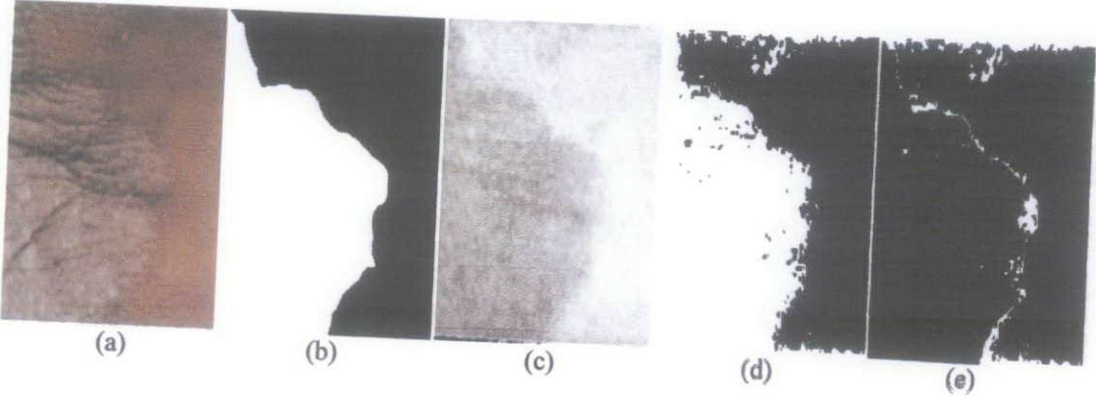
- (a) Original Image, (b) Manual Segmentation, (c) Transformation to  $I_3$  colour space,*
- (d) Thresholded in  $I_3$  space with Otsu's Method,*
- (e) Differences between Manual Segmentation and  $I_3$  space,*
- (f) Transformation to  $b$  colour space, (g) Thresholded in  $b$  space with Otsu's Method,*
- (h) Differences between Manual Segmentation and  $b$  space*

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

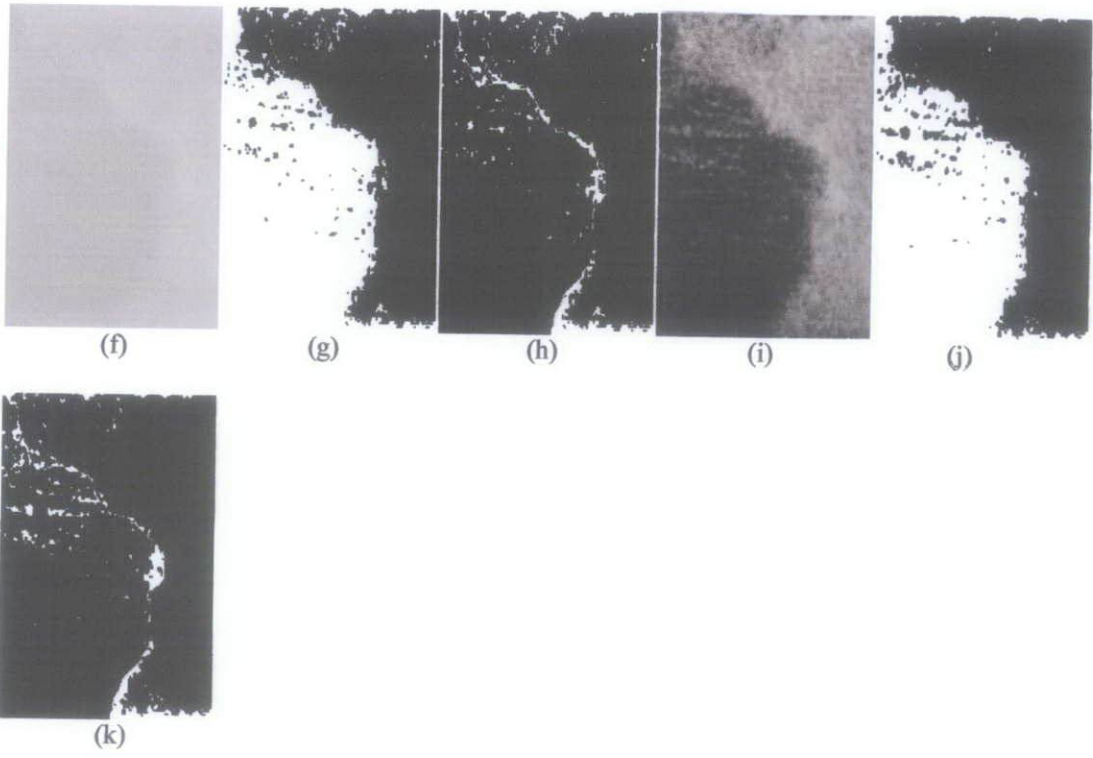
Patch 15



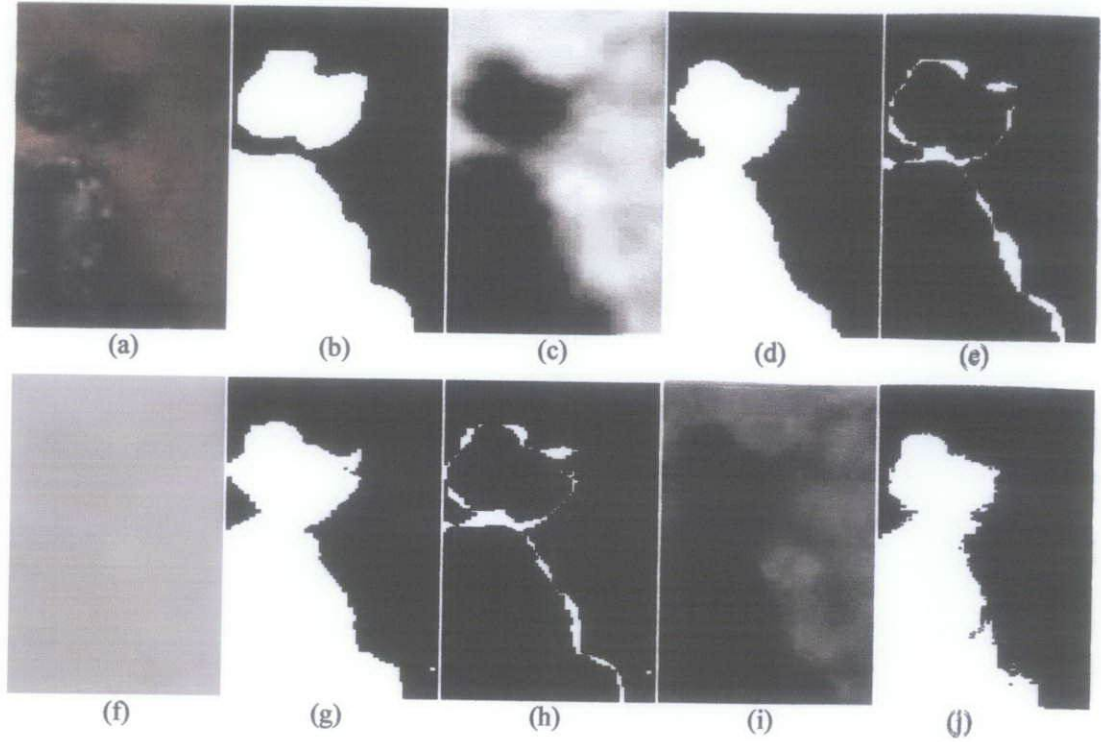
Patch 16







**Patch 17**

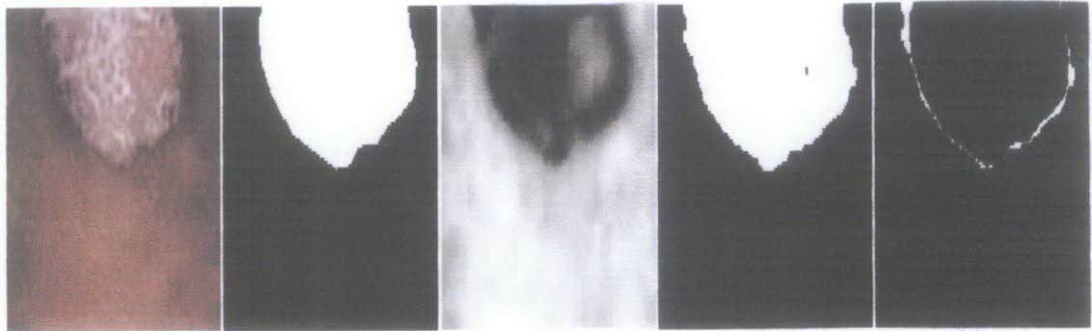






(k)

### Patch 18



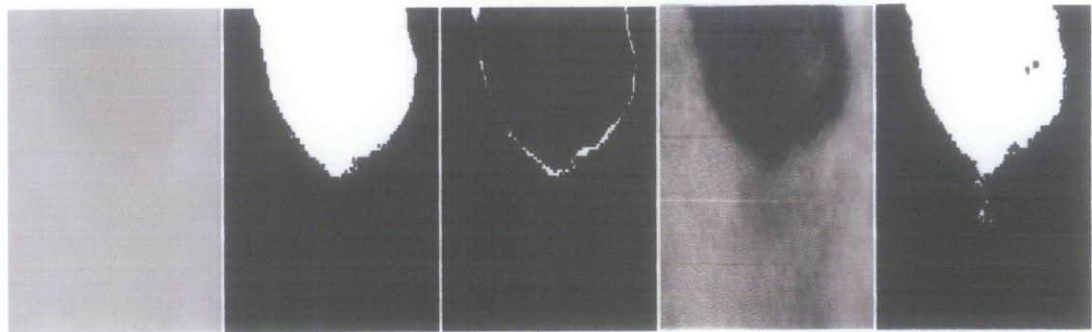
(a)

(b)

(c)

(d)

(e)



(f)

(g)

(h)

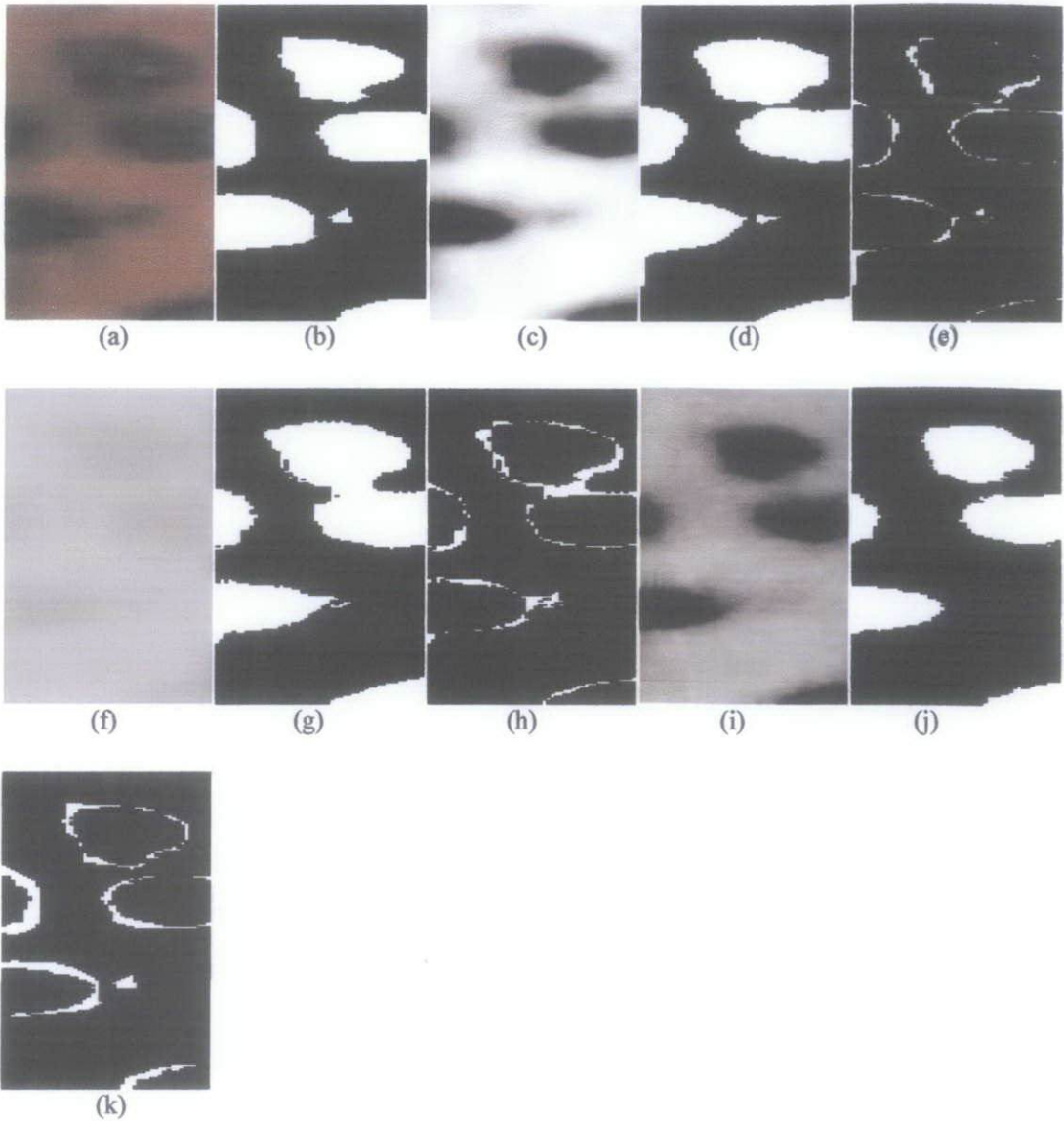
(i)

(j)

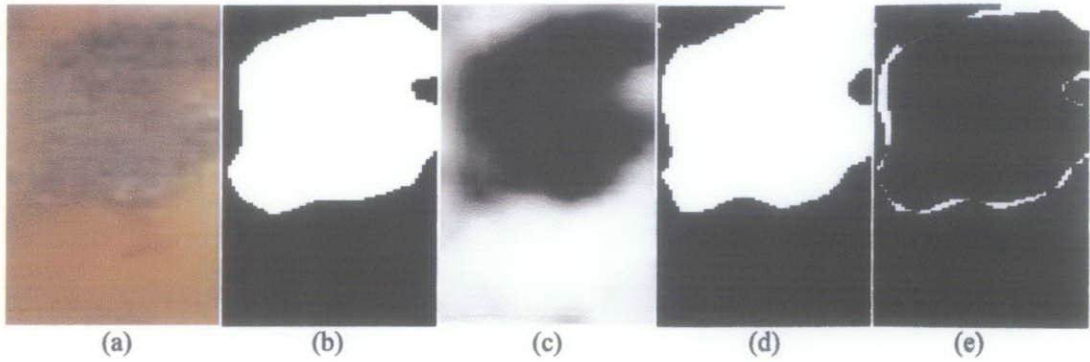


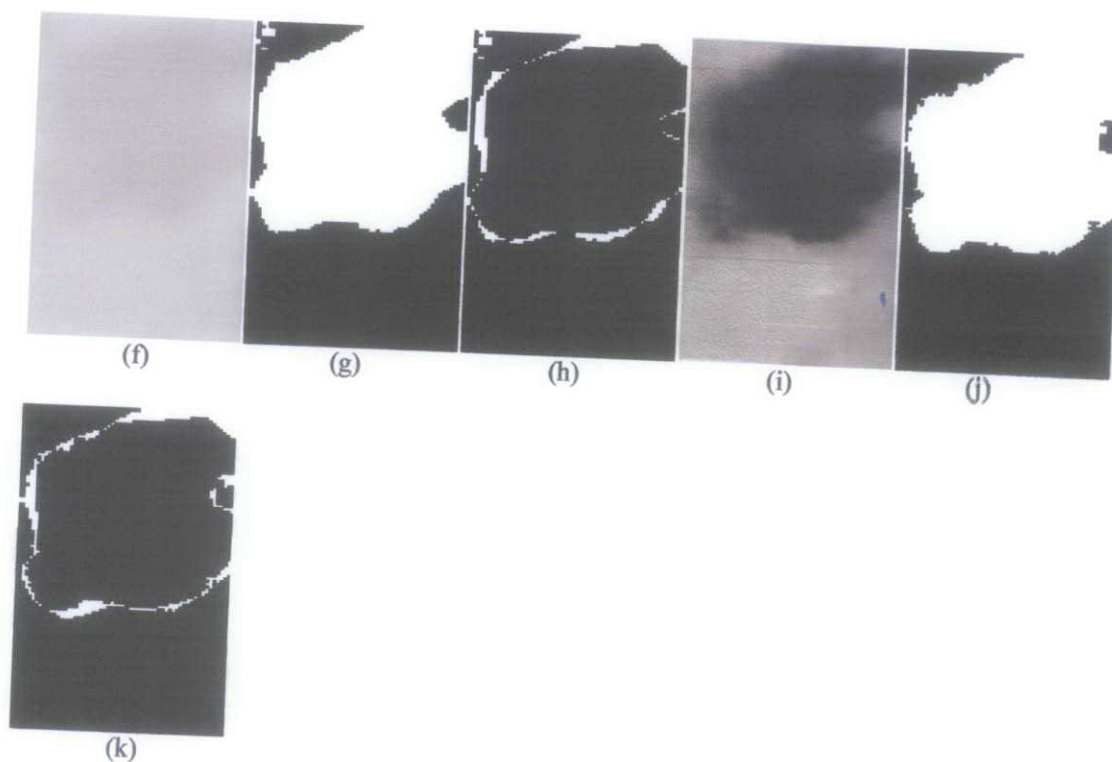
(k)

**Patch 19**



**Patch 20**





**Figure 26: Real Skin-Lesion Patches (Dark Skin Tone) with Appropriate Colour Space Transformation and Otsu's Segmentation (Patch 15 - Patch 20)**

(a) Original Image, (b) Manual Segmentation, (c) Transformation to  $I_2$  colour space, (d) Thresholded in  $I_2$  space with Otsu's Method, (e) Differences between Manual Segmentation and  $I_2$  space, (f) Transformation to  $b$  colour space, (g) Thresholded in  $b$  space with Otsu's Method, (h) Differences between Manual Segmentation and  $b$  space, (i) Transformation to  $S$  colour space, (j) Thresholded in  $S$  space with Otsu's Method, (k) Differences between Manual Segmentation and  $S$  space,

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

**Table 26: Specificity for Segmentation of Skin and Lesion (Patch 1 - Patch 20)**

Skin Lesion Patch #	$I_2$	$I_3$	$b$	$S$
1	-	0.9886	-	-
2	-	0.9184	-	-
3	-	0.9939	-	-
4	-	0.9750	-	-
5	-	0.9741	-	-
6	-	0.9734	-	-
7	-	0.9631	-	-
8	-	0.9385	-	-
9	-	0.9702	0.9702	-
10	-	0.9566	0.9189	-
11	-	0.9196	0.9663	-
12	-	0.9454	0.9631	-

13	-	0.9639	0.9331	-
14	-	0.9295	0.9364	-
15	0.9529	-	0.9546	0.9257
16	0.9681	-	0.9644	0.9378
17	0.9427	-	0.9386	0.9397
18	0.9995	-	0.9800	0.9631
19	0.9395	-	0.9213	0.9152
20	0.9777	-	0.9827	0.9584

*Table 27: Type I error (Patch 1 - Patch 20)*

Skin Lesion Patch #	$I_2$	$I_3$	$b$	$S$
1	-	0.0114	-	-
2	-	0.0816	-	-
3	-	0.0061	-	-
4	-	0.025	-	-
5	-	0.0259	-	-
6	-	0.0266	-	-
7	-	0.0369	-	-
8	-	0.0615	-	-
9	-	0.0298	0.0298	-
10	-	0.0434	0.0811	-
11	-	0.0804	0.0337	-
12	-	0.0546	0.0369	-
13	-	0.0361	0.0669	-
14	-	0.0705	0.0636	-
15	0.0471	-	0.0454	0.0743
16	0.0319	-	0.0356	0.0622
17	0.0573	-	0.0614	0.0603
18	0.0005	-	0.02	0.0369
19	0.0605	-	0.0787	0.0848
20	0.0223	-	0.0173	0.0416

*Table 28: Sensitivity for Segmentation of Skin and Lesion (Patch 1 - Patch 20)*

Skin Lesion Patch #	$I_2$	$I_3$	$b$	$S$
1	-	0.9677	-	-
2	-	0.9939	-	-
3	-	0.9639	-	-
4	-	0.9968	-	-
5	-	0.9563	-	-
6	-	0.9827	-	-
7	-	0.9581	-	-
8	-	0.9715	-	-
9	-	0.9785	0.9785	-

10	-	0.9857	0.9781	-
11	-	0.9687	0.9884	-
12	-	0.9932	0.9906	-
13	-	0.9891	0.9871	-
14	-	0.9982	0.9334	-
15	0.9885	-	0.9795	0.9931
16	0.8833	-	0.8830	0.9319
17	0.9755	-	0.9612	0.9667
18	0.9651	-	0.9853	0.9856
19	0.9533	-	0.9404	0.9944
20	0.9154	-	0.9597	0.9221

*Table 29: Type II error (Patch 1 - Patch 20)*

Skin Lesion Patch #	$I_2$	$I_3$	$b$	$S$
1	-	0.0323	-	-
2	-	0.0061	-	-
3	-	0.0361	-	-
4	-	0.0032	-	-
5	-	0.0437	-	-
6	-	0.0173	-	-
7	-	0.0419	-	-
8	-	0.0285	-	-
9	-	0.0215	0.0215	-
10	-	0.0143	0.0219	-
11	-	0.0313	0.0116	-
12	-	0.0068	0.0094	-
13	-	0.0109	0.0129	-
14	-	0.0018	0.0666	-
15	0.0115	-	0.0205	0.0069
16	0.1167	-	0.117	0.0681
17	0.0245	-	0.0388	0.0333
18	0.0349	-	0.0147	0.0144
19	0.0467	-	0.0596	0.0056
20	0.0846	-	0.0403	0.0779

*Table 30: Percentage of Accuracy for Segmentation of Skin and Lesion (Patch 1 - Patch 20)*

Skin Lesion Patch #	Percentage of Accuracy (Compared with Manually Segmented) / %			
	$I_2$	$I_3$	$b$	$S$
1	-	97.46	-	-
2	-	99.28	-	-
3	-	99.32	-	-
4	-	99.72	-	-
5	-	98.65	-	-

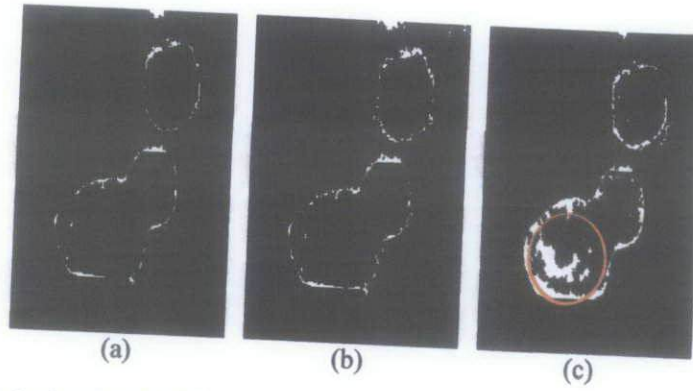


6	-	99.33	-	-
7	-	96.73	-	-
8	-	98.57	-	-
9	-	97.75	97.75	-
10	-	98.91	98.86	-
11	-	98.68	99.10	-
12	-	98.67	98.74	-
13	-	95.62	96.86	-
14	-	99.69	97.78	-
15	99.38	-	99.15	98.09
16	93.49	-	92.69	92.60
17	97.88	-	98.35	96.89
18	99.30	-	99.46	99.17
19	98.31	-	97.82	97.25
20	98.22	-	98.04	98.01

#### 4.4.2 Analysis of Results

Based on the results of *Table 26* and *Table 28*, segmentation on the suggested colour channels has high sensitivity and specificity. Sensitivity relates to the test's ability to identify lesions while specificity relates to the ability of the test to identify skin areas. Type I error shows us the probability of misclassification of lesions as skins and Type II error shows the probability of misclassifications of skins as lesions. Both Type I and Type II errors are low. The results of *Table 30* are analysed. Using Patch #15 on *Figure 30* as a reference, for dark skin-lesion patch, Otsu's thresholding on *S* colour channel is not able to segment edges and the darker parts of the lesion, as indicated by the red circle in *Figure 27(c)*. The saturation values of the darker parts of the lesions are similar to the skin area values. Thus when segmented with Otsu's method, the darker parts of the lesions are misclassified as skin.  $I_2$  and *b* colour channels are more accurate in this respect and have higher segmentation accuracy than *S* (as shown in *Table 30*). The accuracy of segmentation in the *S* colour channel is only approximately 1 % lower than the accuracies in the  $I_2$  and *b* colour channels but in terms of sensitivity, the segmentation results in *S* channel is approximately 3 % lower.



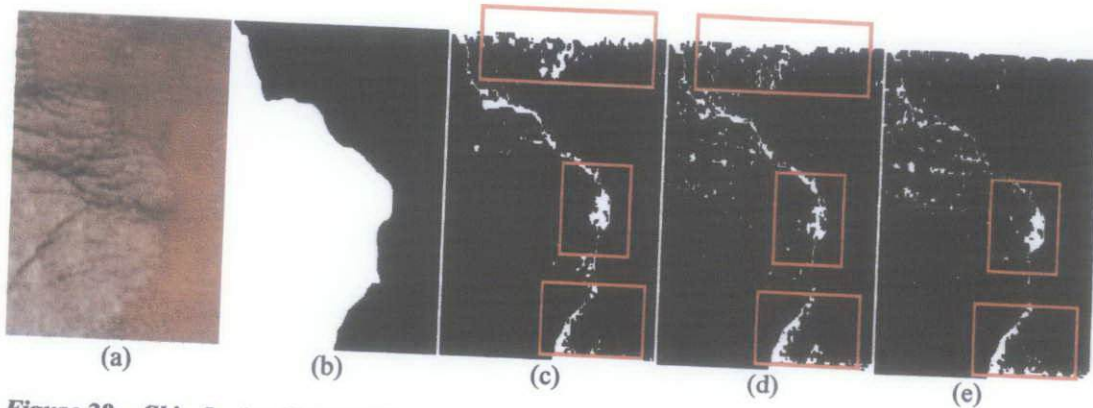


**Figure 27: Skin-Lesion Patch #15**

*(a) Differences between Manual Segmentation and  $I_2$  space, (b) Differences between Manual Segmentation and  $b$  space, (c) Differences between Manual Segmentation and  $S$  space*

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

The accuracies are higher than 95% for segmentation in all colour channels except for the Skin Lesion Patch #16 (indicated by shaded area in *Table 30*). For Patch #16, images of the patch along with the reference segmented images and segmented images are shown in *Figure 28*. It is seen that certain areas of the skin which are very close to the lesion have been segmented incorrectly as lesions, as indicated by the red rectangles in *Figure 28(c), (d) and (e)*.



**Figure 28: Skin-Lesion Patch #16**

*(a) Original Image, (b) Manual Segmentation, (c) Differences between Manual Segmentation and  $I_2$  space, (d) Differences between Manual Segmentation and  $b$  space, (e) Differences between Manual Segmentation and  $S$  space,*

(Source: Images of patients of Dermatology Clinic, Hospital Kuala Lumpur)

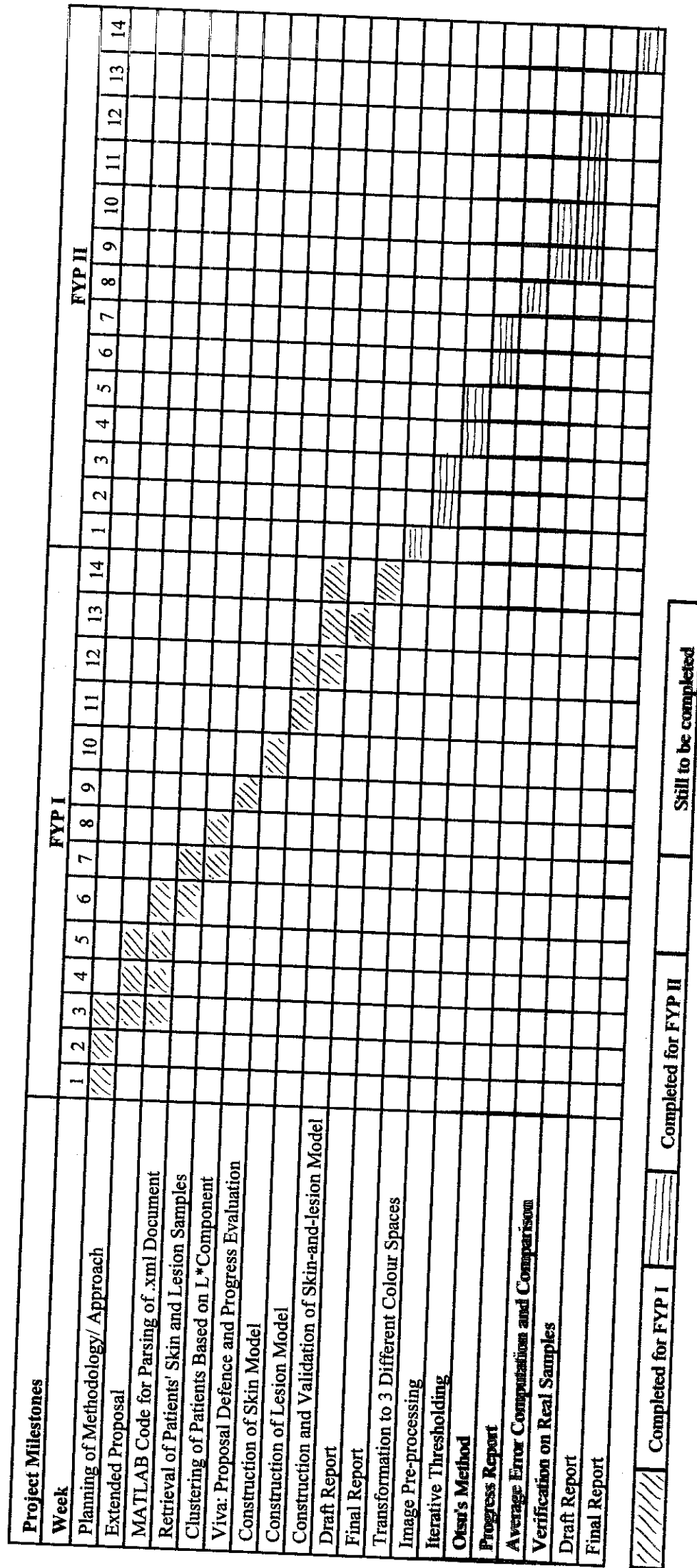
The skin areas in the red rectangles are similar in two aspects: they are wrinkled skin areas and are situated near to the segmented psoriasis lesions. When light is exposed to a smooth patch, light is distributed equally across the entire region. For the

wrinkled skin areas however, the light is not distributed equally. More light tend to concentrate on the peak areas and the trough areas would appear darker. Therefore, these darker regions are misclassified as lesions.

The average error is below 5 % for skin-lesion models in the  $b$  colour channel (middle and dark skin-lesion models),  $I_3$  colour channel (fair and middle skin-lesion models) and  $I_2$  colour channel (dark skin-lesion model). This concurs with the hypotheses of colour spaces advanced in *Chapter 2* as per the studies of *Umbaugh* [14] and X. Yuan, *et al.* [15] for CIE  $L^* a^* b^*$  space and the studies of *Ohta, et al.* [18] and *Dhawan, et al.* [19]. The  $b$  colour channel computes the differences of lightness transformations of (putative) cone responses of the yellow and blue hue, which apparently is a fairly solid distinction between the skin and lesion areas. For the fair and middle skin-lesion models,  $I_3$  provides a good bimodality upon which thresholding technique can be carried out.  $I_3$  represents the deviation from the mean colour (since the red and blue are extreme colours of the colour spectrum and green is considered to be the mean colour). [19] This agrees with the study of *Maletti, et al.* [11] where Wang's Expectation-Maximization Algorithm and fixed discrimination function were carried out. However in this research, Otsu's method works fairly accurate as long as the assumption that the pixels of the classes to be Gaussianly distributed holds, which is true for the  $I_3$  colour channel. When it comes to the dark skin-tone, *Maletti, et al.'s* [11] does not work well since some of the shadows were misclassified as lesions. From *Table 30*, it shows that  $I_2$  colour channel works better than the  $I_3$  colour channel when it comes to the dark skin-lesion model, with approximately 71 % less in average error percentage since  $I_2$  essentially represents the overall variation of colours and is a better representation with better bimodality in the dark skin-lesion model. For the dark skin group, another colour channel that yields high percentage of accuracy when thresholded with the Otsu's method is the  $S$  colour channel. This again tallies with the analysis of the dark skin-lesion models on *page 41* that the bimodality in the  $S$  colour channel is optimal for thresholding.

## **CHAPTER 5: TIMELINE FOR COMPLETION OF FINAL YEAR PROJECT**

**Start Date : 23 May 2011**



## CHAPTER 6: CONCLUSIONS

### 6.1 Discussion

Psoriasis is a widespread skin disease affecting up to 3% of the population. The disease is not curable but there are a number of available treatments to treat psoriasis symptoms. During treatment, dermatologist will monitor the extent of psoriasis continuously to ascertain the treatment efficacy [Pariser, 2003]. Dermatologists usually assess the severity through their tactile sense and thus may not assess the lesion objectively. The Psoriasis Area and Severity Index (PASI) is the gold standard for measuring psoriasis severity index. PASI scoring depends on 4 parameters, namely area, redness, scaliness and thickness [Fredriksson, 1978].

These 4 parameters require the psoriatic lesions to be first segmented from skin samples before the individual scoring can be performed. Preliminary observations of images of lesions of random patients with psoriasis seem to show differences in colour between different lesions even for the same individual (the intra-variability of diagnostic parameters). The difference in colour of lesions (the inter-variability of diagnostic parameters) for differing skin tones too should be taken into account. In view of the aforementioned problems, alternatives such as objective digital image processing algorithms have been written to assess the chromatic differences between lesions belonging to different patients and between lesions belonging to the same individual and subsequently segment the psoriatic lesions as accurately as possible.

This report thus investigates digital image analysis techniques to segment psoriatic lesions. The type of psoriasis lesion which is being analyzed for this work is *Psoriasis Vulgaris* since it accounts for 80% of psoriasis lesions and is the most common. This research investigates iterative thresholding and Otsu's thresholding segmentation on 3 colour spaces, namely,  $I_1I_2I_3$ , CIE  $L^* a^* b^*$  and HSI. The approach is first tried out on 3 manually constructed skin-lesion models corresponding to the 3 different skin tones of dark, brown and fair. The skin-lesion models are constructed with 1000 new colour formed from the combination of R, G and B clusters. In order to select the best colour channel in which to perform the segmentation for either thresholding, the average percentage error is compared.

For the segmentation done on skin-lesion models, the results show that the Otsu's method constantly yields a lower average error percentage as compared to the iterative thresholding method for all instances. This is due to the fact that iterative thresholding when performed on the skin-lesion models is found to terminate its iteration when the change becomes small. This observation tallies with the studies of Riddler [28] and Majid, *et al.* [29]. Based on the segmentation results of skin lesion models, the research focuses on segmentation in the  $I_3$  colour channel (for fair skin tone),  $I_3$  and  $b$  colour channels (for middle skin tone) and  $I_2$ ,  $b$  and  $S$  colour channels (for dark skin tone).

The accuracy performance of the lesion segmentation method is determined by comparing segmented images with reference segmented images. Sensitivity and specificity analysis related to the Type I and Type II errors are performed as well. Reference segmented images are obtained by manually segmenting lesion area from the digital images. The lesion segmentation method is applied on images of 20 skin-lesion patches. Out of 20 cases, the segmentation method achieved accuracies of higher than 95% for 19 cases. The lowest accuracy obtained is for Skin-Lesion Patch #16 with accuracies of 92 - 93%. The lower accuracy is due to some wrinkled skin areas exposed to unequally distributed light leading to misclassification as lesions.

For the other 19 skin-lesion patches, the segmentation results of real skin lesion patches tally with the segmentation results of skin lesion models. Otsu's method produces average error of below 5 % for skin-lesion patches in the  $b$  colour channel (middle and dark skin tones),  $I_3$  colour channel (fair and middle skin tones) and  $I_2$  colour channel (dark skin tone). The  $b$  colour channel computes the differences of lightness transformations of (putative) cone responses of the yellow and blue hue. The results show that the  $b$  colour channel has a good bimodality distinction between the skin and lesion areas for middle and dark skin tones. The  $I_3$  colour channel, which represents the deviation from the mean colour, is a fairly accurate thresholding channel for the fair and middle skin tone. This agrees with the study of Maletti, *et al.* [11]. For the dark skin tone however, the  $I_2$  colour channel, which represents the overall variation of colours, is a better representation with better bimodality instead of the  $I_3$  colour channel. The  $S$  colour channel too yields higher than 95% accuracy for Otsu's method on dark skin lesion patches.



In sum, the work of this research has managed to achieve its objectives of using a digital image analysis technique to accurately segment psoriasis lesions from skin patches. This research employed a low complexity and commonly-used thresholding method, the Otsu's method in different colour channels to segment psoriasis lesions to highly accurate results. Whilst many similar researches have been done, this research would be particularly beneficial for dermatologists since the proposed technique is robust to the differences in colour tone of the skin patches and highly accurate with less than 5% average percentage error.

## **6.2 Contribution & Recommendation**

The main contribution of this research is in the development of a system that is able to objectively segment the psoriasis lesions of patients with various skin tones. The segmentation of lesions is crucial for further assessment of the area, thickness, scaliness and redness of lesions, which form the 4 main parameters of PASI scoring. PASI scoring has the potential to minimize inter- and intra-observation variations.

Many research works have been conducted with the aim to segment and assess psoriasis lesion areas. Certain methods however, suffer from the appearance of shadows. This happens mostly in the darker skin tones and leads to misclassification as lesion areas. [9] Moreover, the methods have not been tested on patients with different skin tones. For this research, a method in segmenting psoriasis lesions based on a digital image analysis technique has been developed. 2 thresholding segmentation techniques are tried out on 3 constructed skin-lesion models, namely dark, middle and fair to select appropriate colour channels that would provide the highest segmentation accuracy. The selected thresholding method and colour channels are applied on patients with various skin colours and the segmentation accuracy was high. Besides, segmentation of darker skin tones in the appropriate colour channel does not misclassify shadowed skin areas as lesions.

This work has investigated the appropriate colour channels to perform segmentation in the 3 different skin tones. To improve the accuracy, the image segmentation technique can be enhanced or filtered with a feature extraction method. Texture has some special characteristics. It cannot be described only by the colour or gray level of a given pixel. Instead, patterns in a neighbourhood region surrounding the pixel

need to be considered. Thus texture analysis is less sensitive to the noise and the change of background brightness in images. [31] This would work particularly well for skin areas which are wrinkled and have thus been misclassified as lesions. Since the texture of lesions and wrinkled skin differs, the feature extraction method can be performed to identify the textural differences. One example of such a method is the Hough transform. This method can be used in addition to the segmentation method to further distinguish lesion and skin.

## REFERENCE

- [1] E. M. Farber and M. L. Nall, "The Natural History of Psoriasis in 5,600 Patients," *Dermatology*, vol. 148, pp. 1-18, 1974.
- [2] G. Kavli, *et al.*, "Psoriasis: familial predisposition and environmental factors," *British Medical Journal (Clinical research ed.)*, vol. 291, pp. 999-1000, October 12, 1985 1985.
- [3] B. Sinniah, *et al.*, "Epidemiology of Psoriasis in Malaysia: A Hospital Based," *Med J Malaysia*, vol. 65, p. 2, 2010
- [4] I. Health Grades. (June 12, 2003 24 June ). *Statistics by Country for Psoriasis*
- [5] A. D. Ormerod, *et al.*, "A comparison of subjective and objective measures of reduction of psoriasis with the use of ultrasound, reflectance colorimetry, computerized video image analysis, and nitric oxide production," *Journal of the American Academy of Dermatology*, vol. 37, pp. 51-57, 1997.
- [6] D. Ihtatho, *et al.*, "Area Assessment of Psoriasis Lesion for PASI Scoring," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 3446-3449.
- [7] J. S. Taur, "Neuro-Fuzzy Approach to the Segmentation of Psoriasis Images," *J. VLSI Signal Process. Syst.*, vol. 35, pp. 19-27, 2003.
- [8] J. Rönning, *et al.*, "Area Assessment of Psoriatic Lesions based on Variable Thresholding and Subimage Classification," *Vision Interface '99*, 1999.
- [9] R. Jailani, *et al.*, "Border segmentation on digitized psoriasis skin lesion images," in *TENCON 2004. 2004 IEEE Region 10 Conference*, 2004, pp. 596-599 Vol. 3.
- [10] D. Delgado, *et al.*, "Automatic Scoring of the Severity of Psoriasis Scaling," *Irish Machine Vision and Image Processing Conference 2004*, p. 6, 2004.

- [11] G. Maletti, *et al.*, "A combined alignment and registration scheme of lesions with psoriasis," *Information Sciences*, vol. 175, pp. 141-159, 2005.
- [12] X. Yuan, *et al.*, "A narrow band graph partitioning method for skin lesion segmentation," *Pattern Recogn.*, vol. 42, pp. 1017-1028, 2009.
- [13] D. C. Tseng and C. H. Chang, "Color segmentation using perceptual attributes," in *Pattern Recognition, 1992. Vol.III. Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on*, 1992, pp. 228-231.
- [14] S. E. Umbaugh, "Computer vision in medicine: color metrics and image segmentation methods for skin cancer diagnosis," Electrical Engineering, University of Missouri/Rolla Rolla, 1990.
- [15] X. Yin, *et al.*, "Hand image segmentation using color and RCE neural network," *Robotics and Autonomous Systems*, vol. 34, pp. 235-250, 2001.
- [16] T. Carron and P. Lambert, "Color edge detector using jointly hue, saturation and intensity," in *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, 1994, pp. 977-981 vol.3.
- [17] H. D. Cheng *et al.*, "Color image segmentation: advances and prospects," *Pattern Recognition*, vol. 34, pp. 2259-2281, 2001.
- [18] Y.-I. Ohta, *et al.*, "Color information for region segmentation," *Computer Graphics and Image Processing*, vol. 13, pp. 222-241, 1980.
- [19] A. P. Dhawan and A. Sim, "Segmentation of images of skin lesions using color and texture information of surface pigmentation," *Computerized Medical Imaging and Graphics*, vol. 16, pp. 163-177, 1992.
- [20] Nevatia, "A Color Edge Detector and Its Use in Scene Segmentation," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 7, pp. 820-826, 1977.
- [21] P. K. Sahoo, *et al.*, "A survey of thresholding techniques," *Comput. Vision Graph. Image Process.*, vol. 41, pp. 233-260, 1988.

- [22] T. Uchiyama and M. A. Arbib, "Color image segmentation using competitive learning," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, pp. 1197-1206, 1994.
- [23] R. Ohlander, *et al.*, "Picture segmentation using a recursive region splitting method," *Computer Graphics and Image Processing*, vol. 8, pp. 313-333, 1978.
- [24] M. Celenk, "A color clustering technique for image segmentation," *Computer Vision, Graphics, and Image Processing*, vol. 52, pp. 145-170, 1990.
- [25] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*: Academic Press, Inc. , 1982.
- [26] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, pp. 62-66, 1979.
- [27] H. Jonsson, "Vision-based Segmentation of Hand Regions for Purpose of Tracking Gestures," Master's In Computer Science, UMEA 2008.
- [28] T. W. Ridler, "Picture Thresholding Using an Iterative Selection Method," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, pp. 630-632, 1978.
- [29] A. Magid, *et al.*, "Comments on Picture thresholding using an iterative selection method," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 20, pp. 1238-1239, 1990.
- [30] Bülent Sankura and M. Sezginb, "Image Thresholding Techniques: A Survey Over Categories," Electric-Electronic Engineering Department, aBoğaziçi University Bebek, İstanbul.
- [31] J. S. Taur, *et al.*, "Segmentation of psoriasis vulgaris images using multiresolution-based orthogonal subspace techniques," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 36, pp. 390-402, 2006.



## APPENDIX

### Calculation of Mean of L\* Component

```
clear all; close all; clc;

cd('F:\FYP docs\data\trial\');
fold=dir;
Lt_total = 0;

for i = 3:92

    path=['F:\FYP docs\data\trial\',fold(i).name,'\skin\'];
    disp(path)
    cd(path);

    fileList = dir('*.jpg');

    k= length(dir) - 2;
    idxImg = 1;

    Lt = [];

    for idxImg = 1:k
        I = imread([path,fileList(idxImg).name]);

        rgbImage = I;

        % Create transformation function from RGB to CIE L*a*b space
        rgbToLab = makecform('srgb2lab');
        LabMat = applycform(rgbImage,rgbToLab);

        % Extract L, a, b components
        L = LabMat(:, :, 1);
        a = LabMat(:, :, 2);
        b = LabMat(:, :, 3);
        [m,n]=size(L);
        Lld=reshape(L,m*n,1);

        %         figure (1),imhist(Lt);

        Lt = [Lt; Lld];

    end

    % Normalize the values for each patient

    %         N = length(Lt);
    %         xHist=zeros(256,1);
    %
    %         for i=1:N
    %             val = Lt(i) + 1;
```

```

%           xHist(val) = xHist(val) + 1;
%       end;
%
%       normX = xHist/N;

%       figure,
%       plot(normX);
%       axis([0 255 0 max(normX)*1.2]);

% Calculate mean of graph
%       title(num2str(mean(Lt)));
%       l(i) = mean(Lt);
%       disp(mean(Lt));

% Stacking of graph for different patients
%       Lt_total = [Lt_total; Lt];

% Normalize the values for all patients

%       N = length(Lt_total);

%       xHist=zeros(256,1);

%       for i=1:N
%           val = Lt_total(i) + 1;
%           xHist(val) = xHist(val) + 1;
%       end;

%       normX = xHist/N;

end

%       figure,
%       plot(normX);
%       axis([0 255 0 max(normX)*1.2]);

```

### **Calculation of Threshold for Clustering**

```

k = length(xHist);
total_pixel = 0;
j1 = 0;
j2 = 0;

for i = 1:k

    total_pixel = xHist(i) + total_pixel;

end

thres1 = total_pixel/3;

for i = 1:k

```

```

j1 = xHist(i) + j1;

if (j1 >= thres1)
    disp(i)
    ds_thres = i;
    thres1 = 1e10;
end

end

thres2 = 2*total_pixel/3;

for i = 1:k

    j2 = xHist(i) + j2;

    if (j2 >= thres2)
        disp(i)
        fs_thres = i;
        thres2 = 1e10;
    end

end

end

```

### **Clustering of Patients into Appropriate Skin Tone Group**

```

clc;
cd('C:\Users\Chi Hung\Desktop\data\');
fold=dir;

for i = 3:92

    path=['C:\Users\Chi Hung\Desktop\data\',fold(i).name];

    if (l(i)<= ds_thres)

        oldname = path;
        newname = 'C:\Users\Chi Hung\Desktop\data\Dark';
        movefile(oldname,newname);

    end

    if (l(i)>= ds_thres && l(i) <= fs_thres)

        oldname = path;
        newname = 'C:\Users\Chi Hung\Desktop\data\Middle';
        movefile(oldname,newname);

    end

    if (l(i)>= fs_thres)

```

```

        oldname = path;
        newname = 'C:\Users\Chi Hung\Desktop\data\Fair';
        movefile(oldname,newname);

    end

end

```

### **Construction of Skin-and-Lesion Model MATLAB code**

```

clear all
close all
clc

y = dir('E:\FYP docs\data\mean\middle\*.mat');

r = [];
l = [];

for i = 1:10
    r(i) = 0;
    lr(i) = 0;
    g(i) = 0;
    lg(i) = 0;
    b(i) = 0;
    lb(i) = 0;
end

% for mat_indx = 1:1

mat_indx = 1;
load(['E:\FYP docs\data\mean\middle\' , y(mat_indx).name]);

    for i = 1:10

        r(i) = sum(C.R(i).data) + r(i);
        lr(i) = length(C.R(i).data) + lr(i);

    end

    for i = 1:10

        g(i) = sum(C.G(i).data) + g(i);
        lg(i) = length(C.G(i).data) + lg(i);

    end

    for i = 1:10

        b(i) = sum(C.B(i).data) + b(i);
        lb(i) = length(C.B(i).data) + lb(i);
    end

```

```

        end

        clust_val_r = [];
        clust_val_g = [];
        clust_val_b = [];

    for i = 1:10

        clust_val_r(i) = r(i)/lr(i);
        clust_val_g(i) = g(i)/lg(i);
        clust_val_b(i) = b(i)/lb(i);

    end

    % clear all
    close all
    clc

    rgb = [];
    i = 1;
    r = [];
    g = [];
    b = [];

    for x = 0:9
        for y = 0:9
            for z = 0:9

                r(i) = x;
                g(i) = y;
                b(i) = z;

                i = i+1;

            end
        end
    end

    r = r';
    g = g';
    b = b';

    rgb = [r g b];

    array=[];
    array(1)=12.25;

    for i = 2:10
        array(i) = 25.5 + array(i-1);
    end

    for i = 1:10
        array(i) = floor(array(i));
    end

```



```

end

for i = 1:1000
    for j = 1:3

        if (rgb(i,j) == 0)
            rgb_a (i, j) = array(1);

        elseif (rgb(i,j) == 1)
            rgb_a (i, j) = array(2);

            elseif (rgb(i,j) == 2)
            rgb_a (i, j) = array(3);

            elseif (rgb(i,j) == 3)
            rgb_a (i, j) = array(4);

            elseif (rgb(i,j) == 4)
            rgb_a (i, j) = array(5);

            elseif (rgb(i,j) == 5)
            rgb_a (i, j) = array(6);

            elseif (rgb(i,j) == 6)
            rgb_a (i, j) = array(7);

            elseif (rgb(i,j) == 7)
            rgb_a (i, j) = array(8);

            elseif (rgb(i,j) == 8)
            rgb_a (i, j) = array(9);

            elseif (rgb(i,j) == 9)
            rgb_a (i, j) = array(10);

        end

    end

end

```

```

for i = 1:1000
    j = 1;

    if (rgb(i,j) == 0)
        rgb_b (i, j) = clust_val_r(1,1);

    elseif (rgb(i,j) == 1)
        rgb_b (i, j) = clust_val_r(1,2);

        elseif (rgb(i,j) == 2)
        rgb_b (i, j) = clust_val_r(1,3);

        elseif (rgb(i,j) == 3)

```

```

    rgb_b (i, j) = clust_val_r(1,4);

    elseif (rgb(i,j) == 4)
    rgb_b (i, j) = clust_val_r(1,5);

    elseif (rgb(i,j) == 5)
    rgb_b (i, j) = clust_val_r(1,6);

    elseif (rgb(i,j) == 6)
    rgb_b (i, j) = clust_val_r(1,7);

    elseif (rgb(i,j) == 7)
    rgb_b (i, j) = clust_val_r(1,8);

    elseif (rgb(i,j) == 8)
    rgb_b (i, j) = clust_val_r(1,9);

    elseif (rgb(i,j) == 9)
    rgb_b (i, j) = clust_val_r(1,10);

end

j = 2;

if (rgb(i,j) == 0)
    rgb_b (i, j) = clust_val_g(1,1);

elseif (rgb(i,j) == 1)
    rgb_b (i, j) = clust_val_g(1,2);

    elseif (rgb(i,j) == 2)
    rgb_b (i, j) = clust_val_g(1,3);

    elseif (rgb(i,j) == 3)
    rgb_b (i, j) = clust_val_g(1,4);

    elseif (rgb(i,j) == 4)
    rgb_b (i, j) = clust_val_g(1,5);

    elseif (rgb(i,j) == 5)
    rgb_b (i, j) = clust_val_g(1,6);

    elseif (rgb(i,j) == 6)
    rgb_b (i, j) = clust_val_g(1,7);

    elseif (rgb(i,j) == 7)
    rgb_b (i, j) = clust_val_g(1,8);

    elseif (rgb(i,j) == 8)
    rgb_b (i, j) = clust_val_g(1,9);

    elseif (rgb(i,j) == 9)
    rgb_b (i, j) = clust_val_g(1,10);

```

```

end

j = 3;

if (rgb(i,j) == 0)
    rgb_b (i, j) = clust_val_b(1,1);

elseif (rgb(i,j) == 1)
    rgb_b (i, j) = clust_val_b(1,2);

    elseif (rgb(i,j) == 2)
    rgb_b (i, j) = clust_val_b(1,3);

    elseif (rgb(i,j) == 3)
    rgb_b (i, j) = clust_val_b(1,4);

    elseif (rgb(i,j) == 4)
    rgb_b (i, j) = clust_val_b(1,5);

    elseif (rgb(i,j) == 5)
    rgb_b (i, j) = clust_val_b(1,6);

    elseif (rgb(i,j) == 6)
    rgb_b (i, j) = clust_val_b(1,7);

    elseif (rgb(i,j) == 7)
    rgb_b (i, j) = clust_val_b(1,8);

    elseif (rgb(i,j) == 8)
    rgb_b (i, j) = clust_val_b(1,9);

    elseif (rgb(i,j) == 9)
    rgb_b (i, j) = clust_val_b(1,10);

end

end

clc

yR = dir('E:\FYP docs\data\xmat_for_middle\middle_red\*.mat');
yG = dir('E:\FYP docs\data\xmat_for_middle\middle_green\*.mat');
yB = dir('E:\FYP docs\data\xmat_for_middle\middle_blue\*.mat');

t = 1;

colour = [];

for i =1:1000
    colour(i)=0;
end

```

```

i = 1;
mat_indx = 1;

% for mat_indx = 1:50

    load(['E:\FYP
docs\data\xmat_for_middle\middle_red\',yR(mat_indx).name]);

    load(['E:\FYP
docs\data\xmat_for_middle\middle_green\',yG(mat_indx).name]);

    load(['E:\FYP
docs\data\xmat_for_middle\middle_blue\',yB(mat_indx).name]);

%     for t = 1:length(array_red)

for t = 1:99999

%     while(key)
    for i = 1:1000

        if (floor(array_red(t))/25.5 == rgb(i,1) &&
floor(array_green(t))/25.5 == rgb(i,2)  && floor(array_blue(t))/25.5
== rgb(i,3))

            colour(i) = colour(i)+1;
            key = 0;
        end
    end

end

%% dark skin
load('E:\FYP docs\data\xmat_for_dark\dark_1000_removed_80%.mat');
dark = [];
dark = colour;
r_dark = [];
g_dark = [];
b_dark = [];

%normalization for dark skin
n1 = sum(dark);
dark_n = [];
for i = 1:length(dark)
    dark_n(i) = dark(i)/n1;
end

dark_n = dark_n*500*500;

for i = 1:length(dark)
    dark_n(i) = round (dark_n(i));
end

```

```

array_dark = [];

for i = 1:1000

    array_dark = [array_dark, i*(ones(1,dark_n(i)))];

end

for l = sum(dark_n):250000
    array_dark(l) = floor(rand(1)*1000);
end
m = randperm(length(array_dark));
m2 = reshape(m,500,500);
k1 = zeros(500,500);
g1 = zeros (1000,1);

for i1 = 1:500
    for j1 = 1:500
        k1(i1,j1) = array_dark(m2(i1,j1));
%         g1(k1(i1,j1)) = g1(k1(i1,j1))+1;
    end
end

i = 1; j = 1;

for i = 1:500
    for j = 1:500

        r_dark(i,j) = rgb_b(k1(i,j),1);
        g_dark(i,j) = rgb_b(k1(i,j),2);
        b_dark(i,j) = rgb_b(k1(i,j),3);

    end
end

A1=uint8(r_dark);
A1(:,:,2) = uint8(g_dark);
A1(:,:,3) = uint8(b_dark);
title ('dark skin model')
imwrite (A1, 'E:\FYP docs\al_data_PC2\Dark_Skin_differentmeans.bmp')

%% middle skin
load('E:\FYP docs\data\xmat_for_middle\middle_1000_removed.mat');
middle = [];
middle = colour;

%normalization for middle skin
n2 = sum(middle);
middle_n = [];
for i = 1:length(middle)
    middle_n(i) = middle(i)/n2;
end

```

```

middle_n = middle_n*500*500;

for i = 1:length(middle)
    middle_n(i) = round (middle_n(i));
end

array_middle = [];

for i = 1:1000

    array_middle = [array_middle, i*(ones(1,middle_n(i)))];

end

for l = sum(middle_n):250000
    array_middle(l) = floor(rand(1)*1000);
end

m3 = randperm(length(array_middle));
m4 = reshape(m3,500,500);
k2 = zeros(500,500);

for i1 = 1:500
    for j1 = 1:500
        k2(i1,j1) = array_middle(m4(i1,j1));
%        g1(k1(i1,j1)) = g1(k1(i1,j1))+1;
    end
end

i = 1; j = 1;

for i = 1:500
    for j = 1:500

        r_middle(i,j) = rgb_b(k2(i,j),1);
        g_middle(i,j) = rgb_b(k2(i,j),2);
        b_middle(i,j) = rgb_b(k2(i,j),3);

    end
end

A2=uint8(r_middle);
A2(:,:,2) = uint8(g_middle);
A2(:,:,3) = uint8(b_middle);
title ('middle skin model')
imwrite (A2, 'E:\FYP
docs\al_data_PC2\Middle_Skin_differentmeans.bmp')

%% fair skin
load('E:\FYP docs\data\xmat_for_fair\20%_removed.mat');
fair = [];
fair = colour;

```



```

%normalization for fair skin
n3 = sum(fair);
fair_n = [];
for i = 1:length(fair)
    fair_n(i) = fair(i)/n3;
end

fair_n = fair_n*500*500;

for i = 1:length(fair)
    fair_n(i) = round (fair_n(i));
end

array_fair = [];

for i = 1:1000

    array_fair = [array_fair, i*(ones(1,fair_n(i)))];

end

for l = sum(fair_n):250000
    array_fair(l) = floor(rand(1)*1000);
end

m5 = randperm(length(array_fair)-1);
m6 = reshape(m5,500,500);
k3 = zeros(500,500);

for il = 1:500
    for jl = 1:500
        k3(il,jl) = array_fair(m6(il,jl));
    %
        gl(kl(il,jl)) = gl(kl(il,jl))+1;
    end
end

i = 1; j = 1;

for i = 1:500
    for j = 1:500

        r_fair(i,j) = rgb_b(k3(i,j),1);
        g_fair(i,j) = rgb_b(k3(i,j),2);
        b_fair(i,j) = rgb_b(k3(i,j),3);

    end
end

A3=uint8(r_fair);
A3(:,:,2) = uint8(g_fair);

```

```

A3(:,:,3) = uint8(b_fair);
title ('fair skin model')
% imwrite (A3, 'E:\FYP
docs\al_data_PC2\Fair_Skin_differentmeans.bmp')

```

## Transformation to $I_1I_2I_3$ MATLAB code

```

clear all;
clc
close all

I =
imread('E:\FYPdocs\lesionsegmentation0\colourBmps\Dark\c_les_17.bmp'
);
imshow(I);

I2 = (1/3)*(I(:,:,1)+I(:,:,2)+I(:,:,3));
% figure(2),imshow(I2);

R = I(:,:,1);
G = I(:,:,2);
B = I(:,:,3);

figure, subplot(221), imshow(I2);
subplot(222), imshow(R);
subplot(223), imshow(G);
subplot(224), imshow(B);

R = double(I(:,:,1));
G = double(I(:,:,2));
B = double(I(:,:,3));
I2 = uint8((R+G+B)./3);

% figure, imshow(I2)
I1 = I2;
% I2 = uint8(R-B);
% I3 = uint8((2*G-R-B)/2);
I2 = (R-B);
I3 = ((2*G-R-B)/4);

figure,
subplot(221), imshow(I);
title ('Original Image Dark Skin')

subplot(222), imshow(I1);
title ('I1')

subplot(223), imshow(I2,[]);
title ('I2')
% imwrite(I2,[],'E:\FYP docs\data\Dark_lesion_selected\I2.bmp')

% subplot(224),

```

```

figure,
imshow(I3, []);
title ('I3')
% imwrite(I3, 'E:\FYP docs\data\Dark_lesion_selected\I3.bmp')

close all

[m,n] = size(I);
n = n/3;

% imwrite(I1, 'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\I1.bmp');

I2min = min(min(I2));
I2max = max(max(I2));

for i = 1:m
    for j = 1:n

        I2(i,j) = 255*((I2(i,j)-I2min)/(I2max-I2min));

    end
end

I2 = uint8(I2);
figure (1), imshow(I2)

% imwrite(I2, 'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\I2.bmp');

I3min = min(min(I3));
I3max = max(max(I3));

for i = 1:m
    for j = 1:n

        I3(i,j) = 255*((I3(i,j)-I3min)/(I3max-I3min));

    end
end

I3 = uint8(I3);
figure (2), imshow(I3)
% imwrite(I3, 'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\I3.bmp');

```

### **Transformation to CIE L\*a\*b\* MATLAB code**

```

close all
clear all
clc

```

```

I = imread('E:\FYP docs\lesion segmentation
0\colourBmps\c_les_12_crop.bmp');
rgbImage = I;

% Create transformation function from RGB to Lab space
rgbToLab = makecform('srgb2lab');
LabMat = applycform(rgbImage,rgbToLab);

% % Extract L,a, b components
L = LabMat(:,:,1);
a = LabMat(:,:,2);
b = LabMat(:,:,3);

figure,
subplot(221), imshow(I);
title ('Original Image Dark Skin')
subplot(222), imshow(L);
title ('L')
% imwrite(L,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\L.bmp');

subplot(223), imshow(a);
title ('a')
% imwrite(a,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\a.bmp');

subplot(224), imshow(b);
title ('b')
% imwrite(b,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\b.bmp');

```

## Transformation to HSI MATLAB code

```

close all
clear all
clc

I = imread('E:\FYP docs\lesion segmentation
0\colourBmps\Dark\15_leg_back.jpg');

HSV1=rgb2hsv(I);
H=HSV1(:,:,1);
S=HSV1(:,:,2);
V=HSV1(:,:,3);

subplot(2,2,1), imshow(H)
% imwrite(H,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\J.bmp');

subplot(2,2,2), imshow(S)

```

```

% imwrite(S,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\J.bmp');

subplot(2,2,3), imshow(V)
% imwrite(V,'E:\FYP
docs\data\Sort_excel_model\Dark\Data_retained\V.bmp');

subplot(2,2,4), imshow(I)

```

### **Otsu's MATLAB code**

```

j = I2; %Choose I1, I2, I3, L, a, b, H, S, V

level = graythresh(j);

BW = im2bw(j,level);
% BW = imcomplement (BW);
figure (4), imshow(BW)

```

### **Iterative Thresholding MATLAB code**

```

close all;

[m,n] = size(I);
n = n/3;

z = I3; %choose I1, I2, I3, L, a, b, H, S, V

I4 = double(z(:));

% STEP 1: Compute mean intensity of image from histogram, set T(1) =
mean(I)

[counts, N] = hist(I4, 256);
i = 1;
mul = cumsum(counts);
T(i) = (sum(N.*counts)) / mul(end);

% STEP 2: compute the sample mean of the data classified by the
above threshold
mu2 = cumsum(counts(N<=T(i)));
MBT = sum(N(N<=T(i)).*counts(N<=T(i)))/mu2(end);

mu3 = cumsum(counts(N>T(i)));
MAT = sum(N(N>T(i)).*counts(N>T(i)))/mu3(end);
i=i+1;

% new T = (MAT+MBT)/2
T(i) = (MAT+MBT)/2;

```

```

% STEP 3 repeat step 2 while T(i)~=T(i-1)
thre = T(i);

while abs(T(i)-T(i-1))>=1
    mu2 = cumsum(counts(N<=T(i)));
    MBT = sum(N(N<=T(i)).*counts(N<=T(i)))/mu2(end);

    mu3 = cumsum(counts(N>T(i)));
    MAT = sum(N(N>T(i)).*counts(N>T(i)))/mu3(end);

    i=i+1;
    T(i) = (MAT+MBT)/2;
    thre = T(i);
end

disp(thre)

for i = 1:m
    for j = 1:n
        if z(i,j) > 66
            I6(i,j) = 1;
        else I6(i,j) = 0;
        end
    end
end

figure(5), imshow(I6)

```

### Calculation of Percentage Difference MATLAB code

```

clear all
close all
clc

Ir = imread('F:\FYP docs\lesion segmentation
0\refBmps\bw_11_trunk_back.bmp');
[m,n] = size(Ir);

I = imread('F:\FYP docs\lesion segmentation
0\colourBmps\Dark\11_trunk_back_b.bmp');

diff = abs(im2bw(Ir,0.5)-I);

imshow(diff)

acc = 1 - sum(sum(diff))/(m*n)

```